Chapter 15

Metric spaces and continuity 2

Introduction

In this chapter we build upon the work we did in Chapter 14 by increasing our stock of standard examples of metric spaces. We do this in two ways: by finding ways of forming new metric spaces from existing ones and by looking at some new examples of metrics. In particular, we introduce important examples of metric spaces such as sequence and function spaces.

We also take some time for further examination of continuity between metric spaces and introduce an important class of continuous functions, the *Lipschitz functions*.

In Section 4 we look at what it could mean to say that a sequence of functions converges and in Section 5 we use these ideas to define a metric on a space that consists of continuous functions.

Chapter guide

Although we introduce several techniques and ideas in this chapter, they build on the foundations we laid in the last chapter. In many ways this chapter may be viewed as an opportunity to consolidate and deepen your understanding of metrics.

The examples we give are fundamental in many areas of mathematics. In particular, the space of functions that we consider in Section 5 occurs frequently (and will be studied again in Book F) – you should pay particular attention to this section.

1 New metrics from old

In the previous chapter, we saw several examples of metrics and metric spaces. In this section, we look at some methods by which we can extend our collection of 'known' metric spaces. In particular, in Subsection 1.1 we look at how we can transfer a metric 'across' into a different space: this allows us to formalise the idea of subspaces for a metric space. In Subsection 1.3 we look at how to define products of metric spaces and see that there are several ways of doing this.

1.1 Transferring a metric

Our first method for creating new metric spaces is to find a way of using functions to transfer metrics from one set to another.

Suppose that we have a set X on which we wish to define a metric, and that we also have a metric space (W, d). We wish to 'transfer' the metric d to X; that is, we wish to define a metric on X in terms of the metric d on W. At first sight, it may seem that what we need to do is to bring the metric d from W to X by defining a function from W to X. Rather surprisingly, this does not work effectively, but a suitable function *from* X *into* W does – the idea is that we are finding a function that *embeds* X inside W and this embedding then determines the distance between points in X.

So suppose that $f: X \to W$ and consider two points, x and y, in X. Since f maps from X to W, there are points u and v in W for which $u = f(x)$ and $v = f(y)$. We use the distance (in W) between u and v to define the distance between x and y. Since $u = f(x)$ and $v = f(y)$, this leads us to define a distance d_f on X by $d_f(x, y) = d(u, v) = d(f(x), f(y))$, as shown in Figure 1.1: our hope is that this gives us a metric. As we will shortly see, for this to work, we require $f: X \to W$ to be a *one–one* function.

Remember that a function f is *one–one* if $x = y$ whenever $f(x) = f(y)$.

We first give the function d_f a name.

Definition 1.1 *Pull-back*

Let $f: X \to W$ be a function and let d be a metric on W. The **pull-back of d by f** is the function $d_f: X \times X \to \mathbb{R}$ given by

$$d_f(x, y) = d(f(x), f(y)), \text{ for each } x, y \in X.$$

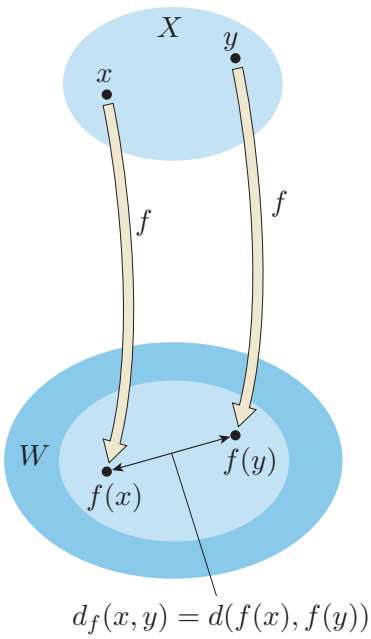


Figure 1.1 Defining $d_f(x, y)$

We now prove that d_f is indeed a metric when f is one–one.

Theorem 1.2

Let $f: X \to W$ be a one–one function and let d be a metric on W. Then the pull-back of d by f, d_f, is a metric on X.

Proof We show that d_f satisfies conditions (M1)–(M3) (Definition 2.1 of Chapter 14).

(M1) The metric d takes only non-negative values. So, for each $x, y \in X$, $d_f(x, y) = d(f(x), f(y)) \geq 0$.

For any $x \in X$, $d_f(x, x) = d(f(x), f(x)) = 0$.

Conversely, if $x, y \in X$ are such that $d_f(x, y) = 0$, then $d(f(x), f(y)) = 0$. Since d is a metric, this implies that $f(x) = f(y)$. But f is one–one, so $x = y$, and d_f satisfies (M1).

This is why f must be one–one.

(M2) For each $x, y \in X$,

$$d_f(y, x) = d(f(y), f(x)) = d(f(x), f(y)) = d_f(x, y),$$

since d is a metric on W. Thus, d_f satisfies (M2).

(M3) Let $x, y, z \in X$. Then, using that d itself satisfies (M3), we find

$$\begin{aligned} d_f(x,z) &= d(f(x), f(z)) \\ &\leq d(f(x), f(y)) + d(f(y), f(z)) \\ &= d_f(x,y) + d_f(y,z). \end{aligned}$$

Thus, d_f satisfies (M3).

Therefore d_f satisfies (M1)–(M3), and so is a metric on X. ■

If f is not one–one, then there are $x, y \in X$ such that $x \neq y$ and $f(x) = f(y)$. For this choice of x and y, we have

$$d_f(x,y) = d(f(x), f(y)) = 0.$$

But this means that d_f fails to satisfy (M1), and so is not a metric. Thus the pull-back of d by f is a metric *if, and only if,* f is one–one.

Worked Exercise 1.3

Show that $d\colon (0,1) \times (0,1) \to \mathbb{R}$ given by $d(x,y) = |y^{-1} - x^{-1}|$ is a metric on the open interval $(0,1)$.

Solution

We could show that d is a metric by verifying directly that it satisfies (M1)–(M3). However, it is easier to observe that if $f\colon (0,1) \to \mathbb{R}$ is given by $f(x) = x^{-1}$, then d is the pull-back of the Euclidean metric $d^{(1)}$ on $\mathbb{R}$ by f, since

$$d(x,y) = |y^{-1} - x^{-1}| = d^{(1)}(f(x), f(y)).$$

Thus, by Theorem 1.2 we need only verify that f is one–one to conclude that d is a metric.

To show that f is one–one, suppose that $x, y \in (0,1)$ and that $f(x) = f(y)$. Then $x^{-1} = y^{-1}$, so that $x = y$. Hence f is one–one, and so d is a metric on $(0,1)$.

Exercise 1.1

Use Theorem 1.2 to show that $d\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$d(x,y) = |\tan^{-1} y - \tan^{-1} x|$$

is a metric on $\mathbb{R}$.

Hint: recall that $\tan^{-1}$ is the inverse function of tan, so that $\tan^{-1} x$ is the unique angle $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ such that $\tan\theta = x$. Its graph is sketched in Figure 1.2.

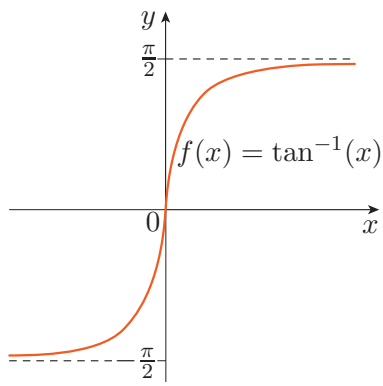


Figure 1.2 Graph of $f(x) = \tan^{-1} x$

Exercise 1.2

Let $d\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by

$$d(x, y) = |y^2 - x^2|.$$

Determine whether d is a metric on $\mathbb{R}$.

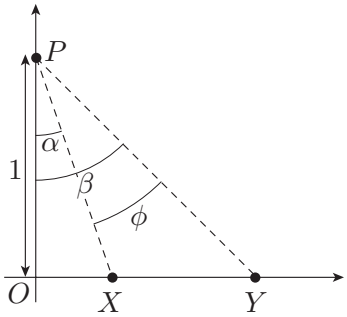


Figure 1.3 Geometric interpretation of the $\tan^{-1}$ metric

The metric defined in Exercise 1.1 is referred to as the **$\tan^{-1}$ metric**, and it has an interesting geometric interpretation.

Consider Figure 1.3. A surveyor at the point P, one unit up the vertical axis, sights the points X and Y (both on the horizontal axis) along the lines PX and PY shown. With P as the vertex, these lines determine an angle ϕ. Let us use ϕ as a measure of the distance d (from the surveyor's perspective) between X and Y.

Let X and Y be at distances x and y respectively from the origin O and assume, without loss of generality, that $y \geq x$: we now find d as a function of x and y. By considering the triangle with vertices P, O and X, we find $\tan\alpha = x$ (since $OP = 1$), so $\alpha = \tan^{-1} x$. Similarly, $\beta = \tan^{-1} y$. Since $y \geq x$, we find that $\phi = \beta - \alpha$ and $d(X, Y) = \phi = \tan^{-1} y - \tan^{-1} x$. In the case $x \geq y$, we would find that $d(X, Y) = \phi = \tan^{-1} x - \tan^{-1} y$. Combining these two cases yields the formula for $d(X, Y)$:

$$d(X, Y) = |\tan^{-1} y - \tan^{-1} x|.$$

Thus, $d(X, Y)$ is the angle between the points X and Y, as observed by the surveyor.

1.2 Metric subspaces

Given a metric on a space X, we often wish to work within a particular subset of our space and use a metric defined on it that 'inherits' properties from the metric on the whole space X. For example, we may be interested in looking at $[0, 1]$ rather than the whole of $\mathbb{R}$ (with the usual Euclidean metric); in such circumstances, we want a metric defined on $[0, 1]$ and not on $\mathbb{R}$ that is consistent with the metric on the whole space.

So suppose we are given a subset A of the underlying set X of a metric space (X, d). We can view A as embedded in X via the function $f\colon A \to X$ given by $f(a) = a$ for each $a \in A$. It is easy to see that this function is one–one and so, from Theorem 1.2, we conclude that the pull-back of d by f, d_f, is a metric on A. But for $a, b \in A$, $d_f(a, b) = d(f(a), f(b)) = d(a, b)$ and so it assigns the right distance between pairs of points in A. We record this in the following definition.

Definition 1.4 *Induced metric*

Let (X, d) be a metric space and let $A \subseteq X$.

The function $d_A \colon A \times A \to \mathbb{R}$ given by

$$d_A(a, b) = d(a, b), \quad \text{for each } a, b \in A,$$

is a metric on A and is called the metric on A **induced by** d, often referred to as the **induced** (or **subspace**) **metric** on A.

The metric space (A, d_A) is called a **metric subspace** of (X, d), and often referred to as the **induced subspace**.

Note that if it is clear from the context, then we may suppress mention of the subset A. For example, we often talk of the Euclidean metric for $[0, 1]$ as $d^{(1)}$ rather than the more cumbersome $d^{(1)}_{[0,1]}$.

Since the metric on a subspace comes from the metric on the original space, one would expect that many of the properties of the subspace will be similar to those of the original space.

Theorem 1.5

Let (X, d) be a metric space and suppose that $A \subseteq X$. Let d_A denote the metric on A induced by d. Then a sequence $(a_n)_{n\in\mathbb{N}}$ in A is d_A-convergent to a if, and only if, $a \in A$ and the sequence (a_n) is d-convergent to a (when viewed as a sequence in X).

Proof Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in A. Then it is also a sequence in X.

Suppose that $a_n \xrightarrow{d_A} a$ as $n \to \infty$. Then, by the definition of convergence, $a \in A$ and $d_A(a_n, a) \to 0$ as $n \to \infty$.

But, by definition, for each $n \in \mathbb{N}$, $d_A(a_n, a) = d(a_n, a)$ and also $a \in A \subseteq X$. Hence $a_n \xrightarrow{d} a$ as $n \to \infty$.

Conversely, suppose that $a \in A$ and the sequence (a_n) d-converges to a as $n \to \infty$.

Then $(d(a_n, a))_{n\in\mathbb{N}}$ is a (real) null sequence and, since $a \in A$ and $a_n \in A$ for each n,

$$d(a_n, a) = d_A(a_n, a) \text{ for each } n \in \mathbb{N}.$$

Hence $(d_A(a_n, a))$ is a (real) null sequence and we conclude that $a_n \xrightarrow{d_A} a$ as $n \to \infty$. ■

Theorem 1.6

Let (X, d) and (Y, e) be metric spaces and suppose that $f\colon X \to Y$ is a function. Let $A \subseteq X$ and let d_A be the metric induced on A by d.

If $a \in A$ and f is (d, e)-continuous at a, then $f|_A$ is (d_A, e)-continuous at a.

(Here the **restriction of f to A**, $f|_A\colon A \to Y$, is given by $f|_A(a) = f(a)$ for each $a \in A$.)

Exercise 1.3

Prove Theorem 1.6.

The idea of an induced metric gives us an easy way to define a metric on common geometric surfaces in $\mathbb{R}^3$, such as the unit sphere with centre $\mathbf{0}$:

$$S = S_{d^{(3)}}(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^3 : d^{(3)}(\mathbf{0}, \mathbf{x}) = 1\}.$$

We simply endow S with the metric induced by the Euclidean metric $d^{(3)}$.

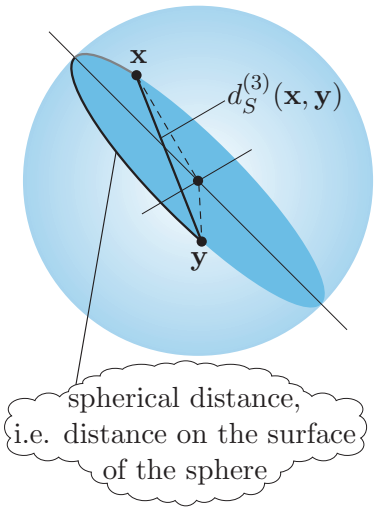


Figure 1.4 Geometric interpretation of $d_S^{(3)}(x, y)$

Exercise 1.4

Show that $(1, 0, 0)$ and $(0, 0, 1)$ are in S, and calculate $d_S^{(3)}((1, 0, 0), (0, 0, 1))$.

Although endowing S with the metric induced by $d^{(3)}$ gives us a metric on the unit sphere, it does not measure distance *on* the sphere in the way that we do on the Earth, as Figure 1.4 illustrates. It is possible to define a metric that does this, but it is a little messy and we omit it.

Exercise 1.5

Let S be the shape given in Figure 1.5, i.e. S consists of three sides of a unit square:

$$S = \{(0, y) : 0 \leq y \leq 1\} \cup \{(x, 0) : 0 \leq x \leq 1\} \cup \{(1, y) : 0 \leq y \leq 1\}.$$

(You could think of S as being a piece of string of length 3 laid around the sides of a square of side 1.)

Let $f\colon S \to \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} 1 - y, & \text{if } x = 0 \text{ and } 0 < y \leq 1, \\ x + 1, & \text{if } y = 0, \\ y + 2, & \text{if } x = 1 \text{ and } 0 < y \leq 1. \end{cases}$$

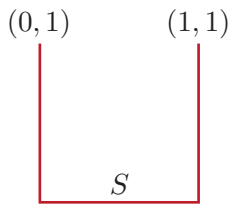


Figure 1.5 S consists of three sides of a unit square: the left side, the base side and the right side

Define a distance function $d\colon S \times S \to \mathbb{R}$ by $d(\mathbf{a}, \mathbf{b}) = |f(\mathbf{a}) - f(\mathbf{b})|$.

(a) Show that f is a one–one function.

(b) Show that d is a metric for S.

(c) For $\mathbf{a} = (0, 1)$ and $\mathbf{b} = (1, 1)$, find $d(\mathbf{a}, \mathbf{b})$ and $d_S^{(2)}(\mathbf{a}, \mathbf{b})$, where $d_S^{(2)}$ is the metric on S induced by the Euclidean metric on $\mathbb{R}^2$.

1.3 Product metrics

We have now seen how to transfer a metric from one set to another (or to a subset of itself) by a one–one function. However, suppose instead that we have a set X defined as the *product* $X_1 \times X_2$ of two sets, each with a metric defined on it. Can we construct a metric on the product set X from the metrics on X_1 and X_2?

$X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1 \text{ and } x_2 \in X_2\}$

You have already come across three examples where this can be done, all involving $X = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, where $X_1 = X_2 = \mathbb{R}$ with the Euclidean metric $d^{(1)}$. The first is the Euclidean metric $d^{(2)}$ on $\mathbb{R}^2$, which can be written in terms of the Euclidean metric $d^{(1)}$ on $\mathbb{R}$ as

$$\begin{aligned} d^{(2)}(\mathbf{x}, \mathbf{y}) &= \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} \\ &= \sqrt{(d^{(1)}(x_1, y_1))^2 + (d^{(1)}(x_2, y_2))^2}. \end{aligned}$$

The second is the taxicab metric on $\mathbb{R}^2$, which can be written in terms of the Euclidean metric $d^{(1)}$ on $\mathbb{R}$ as

$$e_1(\mathbf{x}, \mathbf{y}) = d^{(1)}(x_1, y_1) + d^{(1)}(x_2, y_2).$$

And the third is the 'mixed' metric d on $\mathbb{R}^2$:

$$d(\mathbf{x}, \mathbf{y}) = d^{(1)}(x_1, y_1) + d_0(x_2, y_2) = |y_1 - x_1| + d_0(x_2, y_2),$$

where d_0 denotes the discrete metric for $\mathbb{R}$.

In fact, if (X_1, d_1) and (X_2, d_2) are metric spaces, then there are many metrics that we can construct on the product set $X = X_1 \times X_2$ from the metrics d_1 and d_2. Such metrics are referred to as **product metrics**. We look in detail at three of them.

Theorem 1.7

Let (X_1, d_1) and (X_2, d_2) be metric spaces, and let $X = X_1 \times X_2$. The functions e_1, e_2 and e_∞ from $X \times X$ to $\mathbb{R}$, given by

$$\begin{aligned} e_1(\mathbf{x}, \mathbf{y}) &= d_1(x_1, y_1) + d_2(x_2, y_2), \\ e_2(\mathbf{x}, \mathbf{y}) &= \sqrt{(d_1(x_1, y_1))^2 + (d_2(x_2, y_2))^2}, \\ e_\infty(\mathbf{x}, \mathbf{y}) &= \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}, \end{aligned}$$

for each $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2) \in X$, are metrics on X.

Remarks

1. The subscripts 1 and 2 indicate that e_1 and e_2 are formed by using first and second powers respectively. More generally, if we consider an arbitrary power $p \geq 1$ and define

$$e_p(\mathbf{x}, \mathbf{y}) = (d_1(x_1, y_1)^p + d_2(x_2, y_2)^p)^{1/p},$$

We will not prove this.

then e_p is also a metric on X. Moreover, if we take the limit as $p \to \infty$, then we find that $e_p(\mathbf{x}, \mathbf{y}) \to e_\infty(\mathbf{x}, \mathbf{y})$.

2. In the proof that e_∞ is a metric, we use the following fact:

$$\text{if } a, b, c, d \in \mathbb{R}, \text{ then } \max\{a+b, c+d\} \leq \max\{a, c\} + \max\{b, d\}. \quad (1)$$

This follows easily by considering $a + b$ and $c + d$ separately, and by noting that each of these is at most $\max\{a, c\} + \max\{b, d\}$.

3. The 'mixed' metric that we defined for $\mathbb{R}^2$ in Chapter 14 is formed by using the e_1 construction for the metric spaces $(\mathbb{R}, d^{(1)})$ and $(\mathbb{R}, d_0)$.

You are asked to provide this proof in Exercise 1.6.

Proof The proof that e_1 is a metric is very similar to the proof that the taxicab metric on $\mathbb{R}^2$ is a metric.

The proof that e_2 is a metric is very similar to the proof (given in Chapter 13) that the Euclidean metric for $\mathbb{R}^2$, $d^{(2)}$, satisfies conditions (M1)–(M3).

Here we content ourselves with verifying that e_∞ satisfies conditions (M1)–(M3).

(M1) For each $\mathbf{x}, \mathbf{y} \in X_1 \times X_2$, both $d_1(x_1, y_1)$ and $d_2(x_2, y_2)$ are non-negative, since d_1 and d_2 are metrics, and hence so is their maximum: thus $e_\infty(\mathbf{x}, \mathbf{y}) \geq 0$.

For each $\mathbf{x} \in X_1 \times X_2$,

$$e_\infty(\mathbf{x}, \mathbf{x}) = \max\{d_1(x_1, x_1),\, d_2(x_2, x_2)\} = \max\{0, 0\} = 0,$$

since d_1 and d_2 are metrics.

Conversely, suppose $\mathbf{x}, \mathbf{y} \in X_1 \times X_2$ are such that $e_\infty(\mathbf{x}, \mathbf{y}) = 0$. Then

$$0 = \max\{d_1(x_1, y_1),\, d_2(x_2, y_2)\}.$$

Since d_1 and d_2 are metrics, $d_1(x_1, y_1) \geq 0$ and $d_2(x_2, y_2) \geq 0$, so this equation implies that $d_1(x_1, y_1) = 0$ and $d_2(x_2, y_2) = 0$. Thus, since d_1 and d_2 are metrics, $x_1 = y_1$ and $x_2 = y_2$; that is, $\mathbf{x} = \mathbf{y}$. We conclude that e_∞ satisfies (M1).

(M2) Let $\mathbf{x}, \mathbf{y} \in X_1 \times X_2$. Using property (M2) for the metrics d_1 and d_2 respectively, we have

$$\begin{aligned} e_\infty(\mathbf{y}, \mathbf{x}) &= \max\{d_1(y_1, x_1),\, d_2(y_2, x_2)\} \\ &= \max\{d_1(x_1, y_1),\, d_2(x_2, y_2)\} = e_\infty(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Thus e_∞ satisfies (M2).

(M3) Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X_1 \times X_2$. Then, using property (M3) for the metrics d_1 and d_2,

$$\begin{aligned} e_\infty(\mathbf{x}, \mathbf{z}) &= \max\{d_1(x_1, z_1),\, d_2(x_2, z_2)\} \\ &\leq \max\{d_1(x_1, y_1) + d_1(y_1, z_1),\, d_2(x_2, y_2) + d_2(y_2, z_2)\} \\ &\leq \max\{d_1(x_1, y_1),\, d_2(x_2, y_2)\} + \max\{d_1(y_1, z_1),\, d_2(y_2, z_2)\} \\ &= e_\infty(\mathbf{x}, \mathbf{y}) + e_\infty(\mathbf{y}, \mathbf{z}). \end{aligned}$$

This step is justified by equation (1).

Thus e_∞ satisfies (M3).

Therefore e_∞ is a metric on $X_1 \times X_2$. ■

Exercise 1.6

Let (X_1, d_1) and (X_2, d_2) be metric spaces. Prove that $(X_1 \times X_2, e_1)$ is a metric space.

Product metrics on the plane

When $(X_1, d_1) = (X_2, d_2) = (\mathbb{R}, d^{(1)})$, our methods for constructing metrics on product spaces give us three different ways of defining a metric on $\mathbb{R}^2$:

$$\begin{aligned} e_1(\mathbf{x}, \mathbf{y}) &= d^{(1)}(x_1, y_1) + d^{(1)}(x_2, y_2) = |y_1 - x_1| + |y_2 - x_2|, \\ e_2(\mathbf{x}, \mathbf{y}) &= \sqrt{d^{(1)}(x_1, y_1)^2 + d^{(1)}(x_2, y_2)^2} = \sqrt{|y_1 - x_1|^2 + |y_2 - x_2|^2}, \\ e_\infty(\mathbf{x}, \mathbf{y}) &= \max\{d^{(1)}(x_1, y_1), d^{(1)}(x_2, y_2)\} = \max\{|y_1 - x_1|, |y_2 - x_2|\}. \end{aligned}$$

The metrics e_1 and e_2 are the familiar taxicab and Euclidean metrics on the plane. The metric e_∞ is known as the **max metric** on the plane.

Exercise 1.7

Show that e_1, e_2 and e_∞ are distinct metrics on the plane by calculating $e_1(\mathbf{0}, (2, 1))$, $e_2(\mathbf{0}, (2, 1))$ and $e_\infty(\mathbf{0}, (2, 1))$.

We can obtain an idea of the difference in behaviour between these metrics on the plane by sketching their unit open balls centred at the origin, as shown in Figure 1.6. In Worked Exercise 2.9 of Chapter 14, we saw that for e_1, the taxicab metric, the unit open ball centred at the origin consists of all the points inside a diamond centred at the origin. Since for the plane $e_2 = d^{(2)}$, the Euclidean distance on the plane, we see that for e_2 the unit open ball centred at the origin is an open disc of unit radius centred at the origin. For e_∞, the unit open ball centred at the origin consists of all the points inside the square of side length 2 centred at the origin. In all three cases, since we are considering *open* balls, the boundary points are excluded.

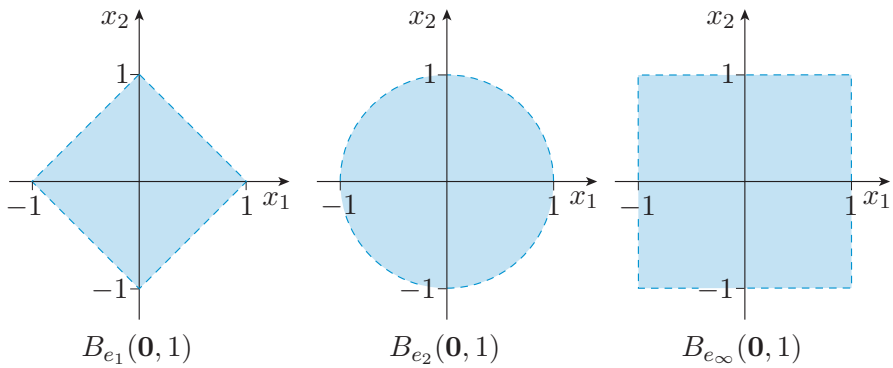


Figure 1.6 Unit open balls centred at the origin for three metrics

Exercise 1.8

(a) Justify the sketch of the unit open ball centred at the origin for e_∞.

(b) What are the unit closed balls and spheres centred at the origin for e_1, e_2 and e_∞?

1.4 n-fold products

Suppose that we are interested in the product of three metric spaces. In principle, this is no more difficult than the product of two: we take the product of the first two, which is a metric space; then we take the product of this new metric space with the third. In doing this, we may use e_1, e_2 or e_∞, or any other product metric that we have defined.

For example, suppose that (X, d) is a metric space; then there are many ways we can define a product metric e on $X^3 = X \times X \times X$. We could take

$$e((a_1, a_2, a_3), (b_1, b_2, b_3)) = d(a_1, b_1) + d(a_2, b_2) + d(a_3, b_3)$$

or

$$e((a_1, a_2, a_3), (b_1, b_2, b_3)) = \sqrt{d(a_1, b_1)^2 + (d(a_2, b_2) + d(a_3, b_3))^2}$$

or many other variants. In all cases, by keeping a clear head and using our results concerning products of pairs of spaces, we can show that these distances e all define metrics on X^3.

This approach can be continued, enabling us to obtain products of n metric spaces, for any given $n \in \mathbb{N}$. We do not go into any great detail here but do take the opportunity to introduce a new example of a metric space in light of this idea.

Example 1.8

Let A be a finite set, called the **set of symbols**, and let d_0 be the discrete metric for A.

Fix $n \in \mathbb{N}$ and let $X = \underbrace{A \times \cdots \times A}_{n} = A^n$, **the words of length n over the set of symbols A.**

Define the Hamming distance $d_H : X \times X \to \mathbb{R}$ for $\mathbf{a}, \mathbf{b} \in X$ by

$$d_H(\mathbf{a}, \mathbf{b}) = \text{the number of places at which } \mathbf{a} \text{ and } \mathbf{b} \text{ differ.}$$

Richard Hamming (1915–1998) was a North American mathematician renowned for his work on error-correcting codes.

It is not too difficult to see that the Hamming distance between pairs of words in X is the product metric e obtained by using e_1, so that for each pair of words $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n) \in X$,

$$e(\mathbf{a}, \mathbf{b}) = d_0(a_1, b_1) + d_0(a_2, b_2) + \cdots + d_0(a_n, b_n).$$

Hence d_H defines a metric on X.

Remarks

1. It is quite common when considering Hamming distance on A^n to write words as strings $a_1 a_2 \cdots a_n$ rather than as vectors $(a_1, a_2, \ldots, a_n)$.
2. A particularly useful case is obtained when $A = \{0, 1\}$ and so A^n consists of all possible binary strings of length n. In this case, the Hamming distance measures in how many places two binary strings of the same length differ – this has applications in the theory of error-correcting codes.

Balls for the Hamming distance

We look briefly at the balls for the Hamming distance on A^n when $A = \{0, 1\}$.

Worked Exercise 1.9

For $n = 2$, find $B_{d_H}((0,0), r)$ for $r > 0$.

Solution

When $n = 2$, $A^n = A^2 = \{(0,0), (1,0), (0,1), (1,1)\}$ and this set of binary strings can be represented as the vertices of a square. Adjacent vertices are distance 1 apart and non-adjacent vertices are distance 2 apart (see Figure 1.7).

Thus

$$B_{d_H}((0,0), r) = \begin{cases} \{(0,0)\}, & \text{if } 0 < r \leq 1, \\ \{(0,0), (1,0), (0,1)\}, & \text{if } 1 < r \leq 2, \\ \{(0,0), (1,0), (0,1), (1,1)\} \, (= A^2), & \text{if } 2 < r. \end{cases}$$

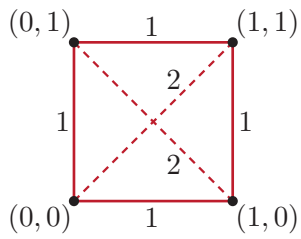


Figure 1.7 Representation of A^2

Exercise 1.9

For $n = 3$, find $B_{d_H}((0,0,0), r)$ for $r > 0$.

Hamming distance can be used as a basis for defining error-correcting codes, as follows. Suppose that we have a string of zeros and ones to transmit. We encode zero as the string 000 and one as the string 111, so the message 010 is transmitted as 000111000. The receiver splits the received message into 'blocks' of length three, and decodes the blocks as 0 if they differ from 000 in no more than one place and 1 if they differ from 111 in no more than one place. It is quite easy to see that following this method, provided no more than one error occurs in each block of length three, the message can still be accurately decoded. Of course, this method of encoding triples the length of each message and there are other, rather more efficient approaches.

2 The Cantor metric

In this section, we give an example of a metric space that is a *sequence space*. Sequence spaces have applications in many areas of mathematics and are often used as a model for understanding the behaviour and structure of more complicated objects, such as *function spaces*.

We discuss an example of a function space later.

Definition 2.1 *Cantor space*

The **Cantor space C** consists of all *infinite* sequences of zeros and ones:

$$\mathbf{C} = \{(a_n) : a_n \in \{0,1\} \text{ for each } n \in \mathbb{N}\}.$$

You can learn more about Georg Cantor (1845–1918) in the *History Reader*.

Remarks

This is a natural extension of the notation for vectors and suggests that one possible way of thinking about points in **C** is to consider them to be infinite vectors.

1. For ease of writing, we denote points in **C** by bold letters: so $\mathbf{x} = (x_n)$ denotes the point $(x_1, x_2, x_3, \ldots)$.
2. Two points in **C** are the same if, and only if, all their terms agree; that is, $\mathbf{x} = \mathbf{y}$ if, and only if, $x_n = y_n$ for each $n \in \mathbb{N}$.
3. There is a useful way of visualising Cantor space in terms of an infinite binary tree, as shown in Figure 2.1.

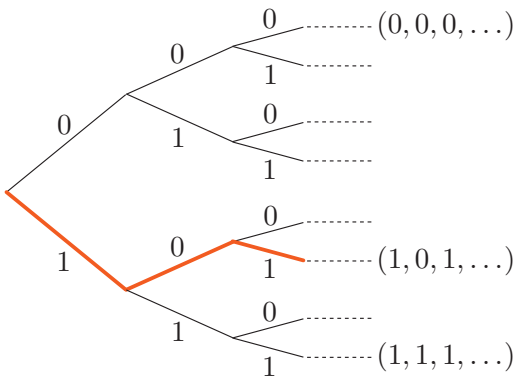


Figure 2.1 An infinite binary tree representing Cantor space

Examples of points in $\mathbf{C}$ are infinite sequences such as

$$(1,1,1,1,\ldots,1,\ldots), \quad (0,0,0,0,\ldots,0,\ldots), \quad (1,0,1,0,\ldots,1,0,\ldots).$$

We have highlighted the start of the path in the tree that represents $(1,0,1,0,\ldots,1,0,\ldots)$.

Suppose we are given two points $\mathbf{x} = (x_n)$ and $\mathbf{y} = (y_n)$ in $\mathbf{C}$. How can we measure the distance between them? We already have one metric that could be used: d_0, the discrete metric. This *would* give us a notion of distance on $\mathbf{C}$ but it would be one that reflected nothing of the structure of $\mathbf{C}$ – we desire a notion of distance that reflects something of the definition of elements of $\mathbf{C}$. One idea would be to follow the concept behind the definition of Hamming distance and count the number of places (or terms) where $\mathbf{x}$ and $\mathbf{y}$ differ. Unfortunately, $\mathbf{x}$ and $\mathbf{y}$ may differ at *infinitely* many terms; thus this definition of distance would have to take infinite values and so would certainly not define a metric.

Consider $(0,0,0,0,\ldots)$ and $(1,1,1,1,\ldots)$, for example.

There are various ways to resolve this problem. The approach we take is to find the first term at which the two sequences differ, and use this as a way of defining the distance between them. Intuitively, such a definition of distance should give smaller values the further 'down' the sequences we must go before we find a term at which the sequences differ.

With this in mind, consider the following four sequences:

$$\mathbf{a} = (0,0,0,0,0,\ldots); \quad \mathbf{b} = (1,1,0,0,0,\ldots);$$
$$\mathbf{c} = (1,0,0,0,0,\ldots); \quad \mathbf{d} = (1,0,1,1,1,\ldots).$$

The sequence $\mathbf{a}$ first differs from $\mathbf{b}$ at the first term; we could say that they are a distance $2^{-1} = \frac{1}{2}$ apart. Next, $\mathbf{b}$ and $\mathbf{c}$ first differ at the second term; we could say that they are a distance $2^{-2} = \frac{1}{4}$ apart. Finally, $\mathbf{c}$ and $\mathbf{d}$ first differ at the third term; we could say that they are a distance $2^{-3} = \frac{1}{8}$ apart. In this way, we get smaller distances the further along the two sequences we have to go before they differ and, of course, if the terms of the two sequences are identical, $x_n = y_n$ for each n, then $\mathbf{x} = \mathbf{y}$.

This leads to the following definition of distance in $\mathbf{C}$.

Definition 2.2 *Cantor distance*

Let $\mathbf{x}, \mathbf{y} \in \mathbf{C}$. The **Cantor distance** $d_{\mathbf{C}} : \mathbf{C} \times \mathbf{C} \to \mathbb{R}$ between $\mathbf{x}$ and $\mathbf{y}$ is given by

$$d_{\mathbf{C}}(\mathbf{x}, \mathbf{y}) = \begin{cases} 0, & \text{if } \mathbf{x} = \mathbf{y}, \\ 2^{-n}, & \text{if } \mathbf{x} \text{ and } \mathbf{y} \text{ first differ at their } n\text{th term.} \end{cases}$$

Note that the Cantor distance between *any* two points of $\mathbf{C}$ is at most $\frac{1}{2}$.

Exercise 2.1

Find $d_{\mathbf{C}}(\mathbf{a}, \mathbf{e})$, $d_{\mathbf{C}}(\mathbf{b}, \mathbf{e})$, $d_{\mathbf{C}}(\mathbf{c}, \mathbf{e})$ and $d_{\mathbf{C}}(\mathbf{d}, \mathbf{e})$ when

$$\mathbf{a} = (1, 0, 0, 0, 0, 0, 0, \ldots), \quad \mathbf{b} = (1, 1, 0, 0, 0, 0, 0, \ldots),$$
$$\mathbf{c} = (1, 1, 1, 0, 0, 0, 0, \ldots), \quad \mathbf{d} = (1, 1, 1, 1, 0, 0, 0, \ldots),$$
$$\mathbf{e} = (1, 1, 1, 1, 1, 1, 1, \ldots).$$

We now show that $d_{\mathbf{C}}$ is a metric on $\mathbf{C}$.

Theorem 2.3

The Cantor distance is a metric on $\mathbf{C}$.

See Chapter 14, Definition 2.1.

Proof We verify that $d_{\mathbf{C}}$ satisfies conditions (M1)–(M3).

(M1) For each $\mathbf{a}, \mathbf{b} \in \mathbf{C}$, we have by definition that $d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) \geq 0$. Moreover $d_{\mathbf{C}}(\mathbf{a}, \mathbf{a}) = 0$.

If $d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 0$ then, by definition, $\mathbf{a} = \mathbf{b}$ and so $d_{\mathbf{C}}$ satisfies (M1).

(M2) Let $\mathbf{a}, \mathbf{b} \in \mathbf{C}$. If $\mathbf{a} = \mathbf{b}$ then, by definition,

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = d_{\mathbf{C}}(\mathbf{b}, \mathbf{a}) = 0.$$

If instead $\mathbf{a} \neq \mathbf{b}$ and $\mathbf{a}$ and $\mathbf{b}$ first differ at the nth term, then $\mathbf{b}$ and $\mathbf{a}$ also first differ at the nth term. Hence,

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = d_{\mathbf{C}}(\mathbf{b}, \mathbf{a}) = 2^{-n}.$$

Thus (M2) holds.

(M3) Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{C}$. We consider four possibilities: (i) $\mathbf{a} = \mathbf{b}$, (ii) $\mathbf{b} = \mathbf{c}$, (iii) $\mathbf{a} = \mathbf{c}$, and (iv) $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are all distinct from each other.

(i) If $\mathbf{a} = \mathbf{b}$, so that $d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 0$, then

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{c}) = d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}) \leq d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) + d_{\mathbf{C}}(\mathbf{b}, \mathbf{c})$$

and so (M3) holds. (In fact $d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}) = d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) + d_{\mathbf{C}}(\mathbf{b}, \mathbf{c})$ here.)

(ii) A similar argument applies if $\mathbf{b} = \mathbf{c}$.

(iii) If $\mathbf{a} = \mathbf{c}$ then $d_{\mathbf{C}}(\mathbf{a}, \mathbf{c}) = 0$, and once again (M3) holds.

(iv) It only remains to consider the case when **a**, **b** and **c** are all distinct. Suppose that **a** first differs from **b** at the mth term and **a** first differs from **c** at the nth term. There are two possible cases: $m \leq n$ and $m > n$.

If $m \leq n$ (see Figure 2.2), then

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{c}) = 2^{-n} \leq 2^{-m} = d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) \leq d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) + d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}), \text{ by (M1)}.$$

If $m > n$ (see Figure 2.3), then **b** and **c** must differ first at the nth term (since **a** agrees with both of them for the first $n-1$ terms and agrees with **b** but not with **c** at the nth term). Thus,

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{c}) = d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}).$$

It follows that

$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{c}) = 2^{-n} = d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}) \leq d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) + d_{\mathbf{C}}(\mathbf{b}, \mathbf{c}), \text{ by (M1)}.$$

Thus, in both cases, (M3) holds.

Since $d_{\mathbf{C}}$ satisfies (M1)–(M3), it is a metric on $\mathbf{C}$. ■

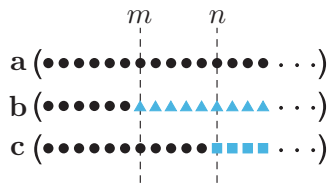


Figure 2.2 Case where $m \leq n$

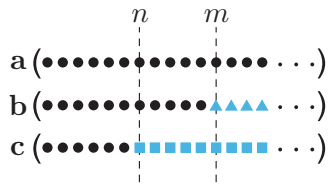


Figure 2.3 Case where $m > n$

Since the Cantor distance is a metric, it is often referred to as the **Cantor metric**.

There are many other examples of sequence spaces in mathematics. Some of the most useful involve sequences of real numbers with specific properties. The method we have given here serves as a model for defining a metric in these more sophisticated settings. In Exercise 2.2, we ask you to show that another sequence space is a metric space.

Exercise 2.2

We say that a sequence of real numbers (a_n) is *eventually zero* if there is an $N \in \mathbb{N}$ such that $a_n = 0$ for each $n > N$. Let ℓ_0 denote the set consisting of all such sequences; that is,

This is a special case of the 'eventually constant' sequence defined in Chapter 14.

$$\ell_0 = \{(a_n) : a_n \in \mathbb{R} \text{ and } (a_n) \text{ is eventually zero}\}.$$

(Notice that different elements of ℓ_0 may have different numbers of non-zero terms.)

Define $d\colon \ell_0 \times \ell_0 \to \mathbb{R}$ by

$$d((a_n), (b_n)) = \sum_{n=1}^{\infty} |b_n - a_n|.$$

This sum is always finite since only finitely many of the terms a_n and b_n are non-zero.

Prove that d defines a metric on ℓ_0.

Balls for the Cantor metric

We now describe open and closed balls for the Cantor metric.

Worked Exercise 2.4

Let $\mathbf{a} = (a_1, a_2, \ldots) \in \mathbf{C}$ and $n \in \mathbb{N}$. Show that for the Cantor metric d,

$$B_d(\mathbf{a}, 2^{-n}) = \{(a_1, a_2, a_3, \ldots, a_n, c_1, c_2, c_3, \ldots) : (c_i) \in \mathbf{C}\}.$$

In other words, the open ball around $\mathbf{a}$ of radius 2^{-n} consists of all infinite sequences of zeros and ones for which the first n terms agree with the first n terms of $\mathbf{a}$ (as shown in Figure 2.4).

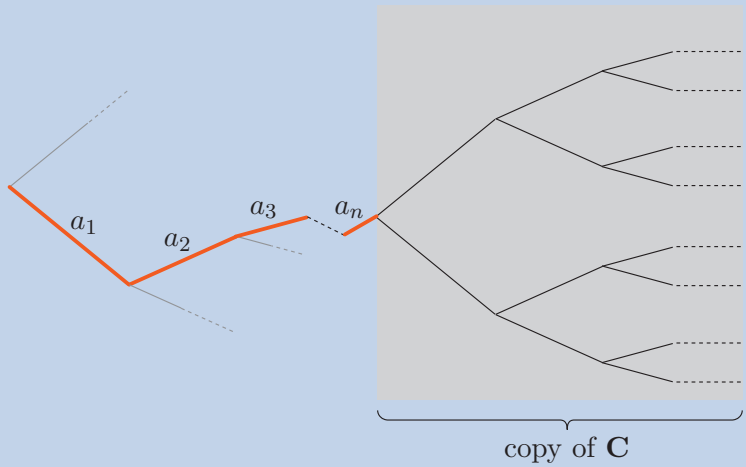


Figure 2.4 The open ball around $\mathbf{a}$ of radius 2^{-n}

Solution

Let $\mathbf{a} = (a_1, a_2, \ldots) \in \mathbf{C}$ and $n \in \mathbb{N}$, and recall that by definition

$$B_d(\mathbf{a}, 2^{-n}) = \{\mathbf{b} \in \mathbf{C} : d(\mathbf{a}, \mathbf{b}) < 2^{-n}\}.$$

Hence if $\mathbf{b} \in \mathbf{C}$ and there is an $i \in \{1, 2, \ldots, n\}$ for which $a_i \neq b_i$, then

$$d(\mathbf{a}, \mathbf{b}) \geq 2^{-i} \geq 2^{-n}$$

and so such a $\mathbf{b}$ is not in $B_d(\mathbf{a}, 2^{-n})$.

Conversely, if $\mathbf{b} \in \mathbf{C}$ agrees with $\mathbf{a}$ for the first n terms, then $\mathbf{b} = (a_1, a_2, \ldots, a_n, b_{n+1}, b_{n+2}, \ldots)$ and $\mathbf{b}$ cannot differ from $\mathbf{a}$ before the $(n+1)$th term. In which case, $d(\mathbf{a}, \mathbf{b}) \leq 2^{-n-1} < 2^{-n}$, implying that $\mathbf{b} \in B_d(\mathbf{a}, 2^{-n})$.

We conclude that

$$B_d(\mathbf{a}, 2^{-n}) = \{(a_1, a_2, a_3, \ldots, a_n, c_1, c_2, c_3, \ldots) : (c_i) \in \mathbf{C}\},$$

as required.

Exercise 2.3

Let $\mathbf{a} = (a_1, a_2, \ldots) \in \mathbf{C}$ and $n \in \mathbb{N}$. Show that for the Cantor metric d,

$$B_d[\mathbf{a}, 2^{-n}] = \begin{cases} \mathbf{C}, & \text{if } n = 1, \\ \{(a_1, a_2, a_3, \ldots, a_{n-1}, c_1, c_2, c_3, \ldots) : (c_i) \in \mathbf{C}\}, & \text{if } n \geq 2. \end{cases}$$

In other words, the closed ball around $\mathbf{a}$ of radius 2^{-n} consists of all infinite sequences of zeros and ones for which the first $n-1$ terms agree with the first $n-1$ terms of $\mathbf{a}$.

Continuity and the Cantor metric

The Cantor space occurs in many different areas of mathematics and there are many examples of continuous functions defined upon it. One simple example is the following.

Definition 2.5 *Shift map*

The **shift map (on C)** is the map $\sigma\colon \mathbf{C} \to \mathbf{C}$ given by

$$\sigma((x_1, x_2, x_3, x_4, \ldots)) = (x_2, x_3, x_4, x_5, \ldots).$$

Thus the shift map acts by deleting the first term of the given sequence. For example,

$$\sigma((0,1,1,1,1,1,\ldots)) = (1,1,1,1,1,\ldots),$$
$$\sigma((0,1,0,1,0,1,\ldots)) = (1,0,1,0,1,\ldots),$$
$$\sigma((1,1,1,1,1,1,\ldots)) = (1,1,1,1,1,\ldots).$$

The shift map 'shifts' the sequence by one term. Such maps play an important role in the theory of dynamical systems.

Worked Exercise 2.6

Show that the shift map is $(d_{\mathbf{C}}, d_{\mathbf{C}})$-continuous on $\mathbf{C}$.

Solution

Let $\mathbf{a} \in \mathbf{C}$ and let $(\mathbf{x}_k)$ be a sequence in $\mathbf{C}$ that converges to $\mathbf{a}$ for $d_{\mathbf{C}}$. We must show that $(\sigma(\mathbf{x}_k))$ converges to $\sigma(\mathbf{a})$ for $d_{\mathbf{C}}$.

To do this, we find an upper bound for $d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a}))$ in terms of $d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a})$. We consider three cases:

1. If $\mathbf{x}_k = \mathbf{a}$, then $d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}) = 0$ and $d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) = 0$.
2. If $\mathbf{x}_k \neq \mathbf{a}$, and $\mathbf{x}_k$ and $\mathbf{a}$ differ at the first term, then $d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}) = \frac{1}{2}$ and $d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) \leq \frac{1}{2}$, since the Cantor distance between two points of $\mathbf{C}$ cannot exceed $\frac{1}{2}$. Thus,
$$d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) \leq \tfrac{1}{2} = d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}).$$

Note that $(\mathbf{x}_k)$ is a sequence of 'points' in the Cantor space $\mathbf{C}$. So for a fixed value of k,
$$\mathbf{x}_k = (x_{k,1}, x_{k,2}, x_{k,3}, \ldots)$$
is a sequence of zeros and ones.

3. If $\mathbf{x}_k \neq \mathbf{a}$, but their first terms are equal, then they must *first* differ at some later term. Suppose that they first differ at the Nth term, where $N \geq 2$. Then $\sigma(\mathbf{x}_k)$ and $\sigma(\mathbf{a})$ first differ at the $(N-1)$th term, and so

$$d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) = 2^{-(N-1)} = 2 \times 2^{-N} = 2d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}).$$

In all three cases,

$$0 \leq d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) \leq 2d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}).$$

Since $(\mathbf{x}_k)$ converges to $\mathbf{a}$ for $d_{\mathbf{C}}$, we have that $d_{\mathbf{C}}(\mathbf{x}_k, \mathbf{a}) \to 0$ as $k \to \infty$. Hence, applying the Squeeze Rule to the inequality above shows that

$$d_{\mathbf{C}}(\sigma(\mathbf{x}_k), \sigma(\mathbf{a})) \to 0 \text{ as } k \to \infty.$$

It follows that $(\sigma(\mathbf{x}_k))$ converges to $\sigma(\mathbf{a})$ for $d_{\mathbf{C}}$, and hence σ is $(d_{\mathbf{C}}, d_{\mathbf{C}})$-continuous at $\mathbf{a}$. Since $\mathbf{a}$ is an arbitrary point of $\mathbf{C}$, we conclude that σ is $(d_{\mathbf{C}}, d_{\mathbf{C}})$-continuous on $\mathbf{C}$.

3 Equivalent metrics

In this section we investigate some conditions under which different metrics on the same underlying set can give rise to the same continuous functions.

3.1 Lipschitz functions

We start by introducing a useful class of functions between metric spaces whose members are always continuous.

Definition 3.1 *Lipschitz function*

Let (X, d) and (Y, e) be metric spaces.

A function $f\colon X \to Y$ is a **Lipschitz function** if there is a non-negative real number M such that, for each $a, b \in X$,

$$e(f(a), f(b)) \leq Md(a, b).$$

The real number M is known as a **Lipschitz constant for** f.

These functions are named after the German mathematician Rudolf Lipschitz (1832–1903).

Remarks

1. When we wish to emphasise the particular metrics d and e on X and Y, we say that f is a (d, e)-*Lipschitz function*.
2. If $(X, d) = (Y, e) = (\mathbb{R}, d^{(1)})$, then the definition reads:

   A function $f\colon \mathbb{R} \to \mathbb{R}$ is a Lipschitz function if there is a non-negative real number M such that, for each $a, b \in \mathbb{R}$,

   $$|f(b) - f(a)| \leq M|b - a|.$$

   By using the Mean Value Theorem it can be shown that $f(x) = \sin x$ satisfies

   $$|\sin b - \sin a| \leq |b - a|$$

   for each $a, b \in \mathbb{R}$; thus the sine function is an example of a $(d^{(1)}, d^{(1)})$-Lipschitz function (with Lipschitz constant $M = 1$).

   See the *Handbook*.
3. The solution of Worked Exercise 2.6 shows that the shift map on $\mathbf{C}$ is a $(d_{\mathbf{C}}, d_{\mathbf{C}})$-Lipschitz function.

The form of the inequality given in the definition is ready-made for proving that Lipschitz functions are continuous.

Theorem 3.2

Let (X, d) and (Y, e) be metric spaces.

If a function $f\colon X \to Y$ is (d, e)-Lipschitz, then it is (d, e)-continuous.

Proof Fix $a \in X$ and let (a_n) be a sequence in X that d-converges to a. We must show that $f(a_n) \to f(a)$ for the metric e as $n \to \infty$.

Since f is Lipschitz, there is a non-negative real number M such that for each $a, b \in X$,

$$e(f(a), f(b)) \leq Md(a, b).$$

In particular, for each $n \in \mathbb{N}$,

$$e(f(a_n), f(a)) \leq Md(a_n, a).$$

Since $a_n \xrightarrow{d} a$ as $n \to \infty$, we know that $(d(a_n, a))$ is a real null sequence, and then the Multiple Rule (for null sequences) gives $Md(a_n, a) \to 0$.

We now apply the Squeeze Rule to deduce that

$$e(f(a_n), f(a)) \to 0 \text{ as } n \to \infty;$$

that is, $(f(a_n))$ e-converges to $f(a)$.

We conclude that f is continuous at a. Since a is an arbitrary point of X, f is continuous on X. ■

We showed that f is continuous in Worked Exercise 2.3 of Chapter 13.

Thus all Lipschitz functions are continuous – but not all continuous functions are Lipschitz. For example, for the Euclidean metric on $\mathbb{R}$, the function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is continuous, but not Lipschitz. For if f were Lipschitz, then there would be a constant $M \geq 0$ such that, for each $a, b \in \mathbb{R}$,

$$|b^2 - a^2| \leq M|b - a|,$$

$|b^2 - a^2| = |b + a|\,|b - a|$.

and so $|b + a| \leq M$ for $a \neq b$. In particular, if $b = 0$ then we would have $|a| \leq M$ for each $a \in \mathbb{R}$, which is clearly false. Hence f is not Lipschitz.

This means that being Lipschitz is more restrictive than being continuous but, despite this, many commonly occurring functions are Lipschitz. In particular, it is often the case that the easiest way to prove that a function is continuous is to prove that it is Lipschitz.

Worked Exercise 3.3

Fix $n \in \mathbb{N}$ and let $X = A^n$ where $A = \{0, 1\}$. Let $f\colon \mathbf{C} \to X$ be given by

$$f((a_1, a_2, a_3, \ldots)) = (a_1, a_2, a_3, \ldots, a_n).$$

(So f maps the infinite binary sequence $\mathbf{a}$ to its first n terms.) Show that f is $(d_{\mathbf{C}}, d_H)$-continuous on $\mathbf{C}$.

Solution

Fix $n \in \mathbb{N}$. We show that f is $(d_{\mathbf{C}}, d_H)$-Lipschitz and use Theorem 3.2.

Let $\mathbf{a} = (a_1, a_2, a_3, \ldots)$ and $\mathbf{b} = (b_1, b_2, b_3, \ldots)$ in $\mathbf{C}$.

We distinguish three cases:

1. If $\mathbf{a} = \mathbf{b}$, then $d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 0$ and $f(\mathbf{a}) = f(\mathbf{b})$ and so, in this case, we find
$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 0 = d_H(f(\mathbf{a}), f(\mathbf{b})).$$
2. If $\mathbf{a} \neq \mathbf{b}$, then suppose they first differ at the mth place. Then
$$d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 2^{-m}.$$
   (a) If $m > n$, then $f(\mathbf{a}) = f(\mathbf{b})$ and so $d_H(f(\mathbf{a}), f(\mathbf{b})) = 0$, giving
   $$0 = d_H(f(\mathbf{a}), f(\mathbf{b})) \leq d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}).$$
   (b) If, instead, $m \leq n$, then it is certainly the case that $d_H(f(\mathbf{a}), f(\mathbf{b})) \leq n$, since the maximum Hamming distance between words of length n is n. But $d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = 2^{-m} \geq 2^{-n}$. Hence if we let $M = n2^n$, a constant that does not depend on the choice of $\mathbf{a}$ and $\mathbf{b}$, then we have
   $$d_H(f(\mathbf{a}), f(\mathbf{b})) \leq n = n\left(2^n \cdot 2^{-n}\right) \leq n2^n \cdot 2^{-m} = Md_{\mathbf{C}}(\mathbf{a}, \mathbf{b}).$$

Thus in all cases we have

$$d_H(f(\mathbf{a}), f(\mathbf{b})) \leq M d_{\mathbf{C}}(\mathbf{a}, \mathbf{b})$$

and so f is $(d_{\mathbf{C}}, d_H)$-Lipschitz (with Lipschitz constant $M = n2^n$). We conclude by Theorem 3.2 that f is $(d_{\mathbf{C}}, d_H)$-continuous on $\mathbf{C}$.

Exercise 3.1

Fix $n \in \mathbb{N}$ and let $X = A^n$ where $A = \{0, 1\}$. Let $f\colon X \to \mathbf{C}$ be given by

$$f((a_1, a_2, \ldots, a_n)) = (a_1, a_2, a_3, \ldots, a_n, 0, 0, 0, \ldots).$$

Show that f is $(d_H, d_{\mathbf{C}})$-continuous on X.

3.2 Equivalence of metrics

We have defined several different notions of product metric for the plane, one of which is the usual definition of Euclidean distance on the plane. This leads to a natural question: if $f\colon \mathbb{R}^2 \to Y$ is a function from $\mathbb{R}^2$ to the metric space (Y, e), does its continuity depend upon which of the three metrics e_1, e_2, e_∞ we used for creating a metric on $\mathbb{R}^2$ (when applied to $d^{(1)}$)? It turns out that it does not: we obtain the same convergent sequences and so, as a result, the same continuous functions. The basic reason why this is the case is that all three metrics are comparable to each other, in that there are simple inequalities that relate any two of them.

We are assuming that each of e_1, e_2 and e_∞ is formed from taking the product of $(\mathbb{R}, d^{(1)})$ with itself – if we change the metric on one of the component spaces for one of the products, there is no reason to suppose that e_1, e_2 and e_∞ will be similar.

Worked Exercise 3.4

Show that for each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, if e_1, e_2 and e_∞ are formed by taking the product of $(\mathbb{R}, d^{(1)})$ with itself, then:

(a) $e_2(\mathbf{x}, \mathbf{y}) \leq e_1(\mathbf{x}, \mathbf{y})$

(b) $e_\infty(\mathbf{x}, \mathbf{y}) \leq e_2(\mathbf{x}, \mathbf{y})$

(c) $e_1(\mathbf{x}, \mathbf{y}) \leq 2e_\infty(\mathbf{x}, \mathbf{y})$.

Solution

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.

(a) We use the fact that if a and b are non-negative real numbers, then $a^2 + b^2 \leq a^2 + 2ab + b^2 = (a + b)^2$, and so $\sqrt{a^2 + b^2} \leq a + b$.

$$\begin{aligned} e_2(\mathbf{x}, \mathbf{y}) &= \sqrt{|y_1 - x_1|^2 + |y_2 - x_2|^2} \\ &\leq |y_1 - x_1| + |y_2 - x_2| = e_1(\mathbf{x}, \mathbf{y}). \end{aligned}$$

(b) $e_\infty(\mathbf{x},\mathbf{y}) = \max\{|y_1 - x_1|,\ |y_2 - x_2|\}$
$\leq \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} = e_2(\mathbf{x},\mathbf{y})$.

(c) $e_1(\mathbf{x},\mathbf{y}) = |y_1 - x_1| + |y_2 - x_2|$
$\leq \max\{|y_1 - x_1|, |y_2 - x_2|\} + \max\{|y_1 - x_1|, |y_2 - x_2|\}$
$= 2\max\{|y_1 - x_1|, |y_2 - x_2|\} = 2e_\infty(\mathbf{x},\mathbf{y})$.

These inequalities imply that for each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$,

$$e_2(\mathbf{x},\mathbf{y}) \leq e_1(\mathbf{x},\mathbf{y}) \leq 2e_2(\mathbf{x},\mathbf{y}), \tag{1}$$
$$e_\infty(\mathbf{x},\mathbf{y}) \leq e_2(\mathbf{x},\mathbf{y}) \leq 2e_\infty(\mathbf{x},\mathbf{y}), \tag{2}$$
$$\tfrac{1}{2}e_1(\mathbf{x},\mathbf{y}) \leq e_\infty(\mathbf{x},\mathbf{y}) \leq e_1(\mathbf{x},\mathbf{y}). \tag{3}$$

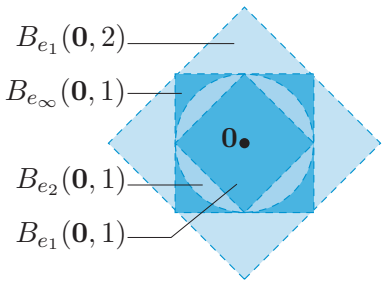


Figure 3.1 Unit open balls for e_1, e_2 and e_∞

One way to understand these inequalities is to consider the relationships between the unit open balls for each of the metrics e_1, e_2 and e_∞. Figure 3.1 shows that $B_{e_1}(\mathbf{0}, 1)$ fits inside $B_{e_2}(\mathbf{0}, 1)$, which fits inside $B_{e_\infty}(\mathbf{0}, 1)$. Moreover, $B_{e_\infty}(\mathbf{0}, 1)$ fits inside a scaled version of $B_{e_1}(\mathbf{0}, 1)$, namely the ball $B_{e_1}(\mathbf{0}, 2)$. By translating and rescaling, we see that versions of these inclusions hold for open balls with any given centre $\mathbf{a}$ and radius r.

Metrics that are related in this way are said to be *(metrically) equivalent*.

Definition 3.5 *Equivalent metrics*

Let d_1 and d_2 be two metrics defined on a set X. The metrics d_1 and d_2 are (metrically) **equivalent** if there are positive real numbers m and M such that for each $x, y \in X$,

$$md_1(x,y) \leq d_2(x,y) \leq Md_1(x,y). \tag{4}$$

Remarks

1. Notice that the inequalities making up equation (4) can be rearranged to give

   $$\tfrac{1}{M}d_2(x,y) \leq d_1(x,y) \leq \tfrac{1}{m}d_2(x,y)$$

   and so there is no asymmetry in the definition: saying that d_1 is (metrically) equivalent to d_2 is the same as saying that d_2 is (metrically) equivalent to d_1.
2. Inequalities (1)–(3) indicate that the metrics e_1, e_2 and e_∞ on $\mathbb{R}^2$ are all equivalent to each other.
3. If we fix the set X and consider the set of all possible metrics for X, then 'metrical equivalence' gives an equivalence relation on the set of metrics for X.

Exercise 3.2

Show that e_1 is equivalent on $\mathbb{R}^2$ to the metric e for $\mathbb{R}^2$ given by

$$e(\mathbf{x}, \mathbf{y}) = |y_1 - x_1| + 2|y_2 - x_2| \text{ for } \mathbf{x} = (x_1, x_2),\ \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2.$$

You may assume that e is a metric on the plane.

The argument used to obtain inequalities (1)–(3) works not only for the product metrics on $\mathbb{R}^2$ but also for the metrics e_1, e_2 and e_∞ defined for a general product $X_1 \times X_2$. We omit the details, which are not very different from those for $\mathbb{R}^2$, and simply record the result.

Theorem 3.6

Let (X_1, d_1) and (X_2, d_2) be metric spaces, and let $X = X_1 \times X_2$. Then the three product metrics e_1, e_2 and e_∞ formed from d_1 and d_2 are equivalent.

We now explore why the notion of equivalence is useful: equivalent metrics have the same convergent sequences.

Proposition 3.7

Let d_1 and d_2 be equivalent metrics defined on the set X. Then a sequence (x_n) converges with limit x in (X, d_1) if, and only if, the same sequence (x_n) converges with the same limit x in (X, d_2).

Proof If the sequence (x_n) d_1-converges to x in X then $(d_1(x_n, x))$ is a real null sequence.

Since d_1 and d_2 are equivalent metrics, there is a fixed positive real number M such that for each $n \in \mathbb{N}$, $0 \leq d_2(x_n, x) \leq M d_1(x_n, x)$.

Thus $(d_2(x_n, x))$ is also a real null sequence by the Squeeze Rule and so (x_n) d_2-converges to x in X.

Conversely, if the sequence (x_n) d_2-converges to a limit x in X, then $(d_2(x_n, x))$ is a null sequence.

Since d_1 and d_2 are equivalent metrics, there is a fixed positive real number m such that for each $n \in \mathbb{N}$, $0 \leq d_1(x_n, x) \leq \frac{1}{m} d_2(x, y)$.

Thus $(d_1(x_n, x))$ is a real null sequence by the Squeeze Rule and so (x_n) d_1-converges with limit x in X. ■

This result has an important corollary.

Theorem 3.8

Let (Y, e) be a metric space and let d_1 and d_2 be metrically equivalent metrics on a set X.

Then

(a) a function $f\colon X \to Y$ is (d_1, e)-continuous at a point $x \in X$ if, and only if, f is (d_2, e)-continuous at x; and

(b) a function $g\colon Y \to X$ is (e, d_1)-continuous at a point $y \in Y$ if, and only if, g is (e, d_2)-continuous at y.

Exercise 3.3

(a) Prove the first part of Theorem 3.8.

(b) Prove the second part of Theorem 3.8.

Since Theorem 3.6 tells us that the metrics e_1, e_2 and e_∞ are equivalent on any product $X_1 \times X_2$, the following result follows immediately.

Theorem 3.9

Let (X_1, d_1), (X_2, d_2) and (Y, e) be metric spaces, let $X = X_1 \times X_2$ and let e_1, e_2 and e_∞ be the three different product metrics on X (determined by d_1 and d_2).

If $f\colon X \to Y$ then, for $j, k \in \{1, 2, \infty\}$, f is (e_j, e)-continuous at $x \in X$ if, and only if, f is (e_k, e)-continuous at x.

If $g\colon Y \to X$ then, for $j, k \in \{1, 2, \infty\}$, g is (e, e_j)-continuous at $y \in Y$ if, and only if, g is (e, e_k)-continuous at y.

Remarks

1. Suppose we have a function f from $X_1 \times X_2$ with one of the product metrics e_1, e_2 or e_∞. Theorem 3.9 tells us that we can test f for continuity by using any one of e_1, e_2 or e_∞, and so in practice we choose the one that is simplest to calculate with.
2. There is a natural extension of this result to the case when X is an n-fold product of metric spaces $X_1, X_2 \ldots, X_n$; we omit the details.

Projection functions and continuity

In Subsection 1.4 of Chapter 14, we saw that we can show a function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is continuous by showing that both $p_1 \circ f\colon \mathbb{R}^2 \to \mathbb{R}$ and $p_2 \circ f\colon \mathbb{R}^2 \to \mathbb{R}$ are continuous, where $p_1\colon \mathbb{R}^2 \to \mathbb{R}$ and $p_2\colon \mathbb{R}^2 \to \mathbb{R}$ are the projection functions defined by $p_1(x_1, x_2) = x_1$ and $p_2(x_1, x_2) = x_2$. We can extend the definitions of projection functions to a general product space, and a similar result about continuity holds.

Definition 3.10 *Projection functions on a space*

Let X_1 and X_2 be sets and let $X = X_1 \times X_2$. The **projection functions** are the functions $p_1\colon X \to X_1$ and $p_2\colon X \to X_2$ given by

$$p_1(x_1, x_2) = x_1 \text{ and } p_2(x_1, x_2) = x_2.$$

As with the projection functions in Chapter 13, these functions are continuous.

Theorem 3.11

Let (X_1, d_1) and (X_2, d_2) be metric spaces and let e be one of the product metrics e_1, e_2 or e_∞ for $X_1 \times X_2$.

Then p_1 is (e, d_1)-continuous and p_2 is (e, d_2)-continuous, where p_1 and p_2 are the projection functions from X to X_1 and X_2 respectively.

Proof Once we have proved the result for any one of e_1, e_2 or e_∞, the others follow immediately from Theorem 3.9.

We prove the result for e_∞ as this turns out to be the simplest to work with.

We show that p_1 is (e_∞, d_1)-continuous; the proof that p_2 is (e_∞, d_2)-continuous is similar.

Fix $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in X$. We find an upper bound on $d_1(p_1(\mathbf{a}), p_1(\mathbf{b}))$ in terms of $e_\infty(\mathbf{a}, \mathbf{b})$:

$$\begin{aligned} d_1(p_1(\mathbf{a}), p_1(\mathbf{b})) &= d_1(a_1, b_1) \\ &\leq \max\{d_1(a_1, b_1), d_2(a_2, b_2)\} = e_\infty(\mathbf{a}, \mathbf{b}). \end{aligned}$$

Thus p_1 is an (e_∞, d_1)-Lipschitz function and so, by Theorem 3.2, p_1 is (e_∞, d_1)-continuous on X. ■

We can now derive the promised continuity result. Since its proof is similar to that of Proposition 1.19 of Chapter 14, we omit it.

Theorem 3.12

Let (Y, d), (X_1, d_1) and (X_2, d_2) be metric spaces, and let (X, e) be the metric space where $X = X_1 \times X_2$ and e is one of the three product metrics e_1, e_2 or e_∞ formed from d_1 and d_2.

Then a function $f: Y \to X_1 \times X_2$ is (d, e)-continuous at $a \in Y$ if, and only if, $p_1 \circ f: Y \to X_1$ is (d, d_1)-continuous at a and $p_2 \circ f: Y \to X_2$ is (d, d_2)-continuous at a.

Theorems 3.11 and 3.12 extend in a natural way to n-fold products of metric spaces $X_1 \times \cdots \times X_n$ once one defines appropriate projection functions $p_i: X_1 \times \cdots \times X_n \to X_i$ by $p_i(x_1, x_2, \ldots, x_n) = x_i$. We omit the details.

Worked Exercise 3.13

Consider the metric spaces $(\mathbf{C}, d_\mathbf{C})$ and $(\mathbb{R}, d^{(1)})$. Define the function $F: \mathbf{C} \to \mathbb{R} \times \mathbf{C}$ by

$$F(\mathbf{a}) = \left(\sum_{n=1}^{\infty} 2^{-n} a_n, \mathbf{a} \right).$$

Notice that the sum in the first component of the definition of $F(\mathbf{a})$ always exists and is finite since each term a_n is either zero or one.

Show that F is $(d_\mathbf{C}, e)$-continuous, where e denotes any of e_1, e_2 or e_∞.

Solution

We consider the composites of the projection functions with F, that is, $p_1 \circ F: \mathbf{C} \to \mathbb{R}$ and $p_2 \circ F: \mathbf{C} \to \mathbf{C}$.

Now, for $\mathbf{a} = (a_1, a_2, a_3, \ldots) \in \mathbf{C}$,

$$(p_1 \circ F)(\mathbf{a}) = \sum_{n=1}^{\infty} 2^{-n} a_n.$$

We show that $p_1 \circ F$ is $(d_\mathbf{C}, d^{(1)})$-Lipschitz. So suppose that $\mathbf{a} = (a_1, a_2, a_3, \ldots), \mathbf{b} = (b_1, b_2, b_3, \ldots) \in \mathbf{C}$. If they are identical, then

$$d^{(1)}((p_1 \circ F)(\mathbf{a}), (p_1 \circ F)(\mathbf{b})) = 0 = d_\mathbf{C}(\mathbf{a}, \mathbf{b}).$$

So suppose instead that they first differ at the mth place. Then $d_\mathbf{C}(\mathbf{a}, \mathbf{b}) = 2^{-m}$ and

$$
\begin{aligned}
&d^{(1)}((p_1 \circ F)(\mathbf{a}), (p_1 \circ F)(\mathbf{b})) \\
&\qquad = \left| \sum_{n=1}^{\infty} 2^{-n}(a_n - b_n) \right| \\
&\qquad \leq \left| \sum_{n=m}^{\infty} 2^{-n}(a_n - b_n) \right|, \text{ since } a_n = b_n \text{ for } n < m \\
&\qquad \leq 2^{1-m} \\
&\qquad = 2d_{\mathbf{C}}(\mathbf{a}, \mathbf{b}).
\end{aligned}
$$

Hence $p_1 \circ F$ is $(d_{\mathbf{C}}, d^{(1)})$-Lipschitz (with Lipschitz constant 2), and so by Theorem 3.2 it is $(d_{\mathbf{C}}, d^{(1)})$-continuous.

Also,

$$(p_2 \circ F)(\mathbf{a}) = \mathbf{a},$$

and so

$$d_{\mathbf{C}}((p_2 \circ F)(\mathbf{a}), (p_2 \circ F)(\mathbf{b})) = d_{\mathbf{C}}(\mathbf{a}, \mathbf{b})$$

for each $\mathbf{a}, \mathbf{b} \in \mathbf{C}$. Thus $p_2 \circ F$ is a $(d_{\mathbf{C}}, d_{\mathbf{C}})$-Lipschitz function (with Lipschitz constant 1), and so by Theorem 3.2 it is $(d_{\mathbf{C}}, d_{\mathbf{C}})$-continuous.

Since both $p_1 \circ F$ and $p_2 \circ F$ are continuous, Theorem 3.12 implies that F is $(d_{\mathbf{C}}, e)$-continuous, where e is any of e_1, e_2 or e_∞.

Exercise 3.4

Consider the metric spaces $(\mathbf{C}, d_{\mathbf{C}})$ and $(\mathbb{R}, d^{(1)})$. Define the function $f\colon \mathbf{C} \to \mathbb{R} \times \mathbf{C}$ by

$$f(\mathbf{a}) = (a_1, \sigma(\mathbf{a})).$$

(a) Evaluate $f(\mathbf{a})$ when:

(i) $\mathbf{a} = (0, 0, 0, 0, 0, 0, \ldots)$

(ii) $\mathbf{a} = (0, 1, 0, 1, 0, 1, \ldots)$.

(b) Show that f is $(d_{\mathbf{C}}, e)$-continuous, where e denotes any of e_1, e_2 or e_∞.

Hint: you may assume that $g\colon \mathbf{C} \to \mathbb{R}$ given by $g(\mathbf{a}) = a_1$ is $(d_{\mathbf{C}}, d^{(1)})$-continuous.

3.3 Convergence in product spaces

We finish this section by taking a brief look at the relationship between convergence of sequences in product spaces and convergence of the corresponding sequences in the component spaces that make up the product.

Let (X, d_1) and (Y, d_2) be metric spaces and let e be one of the product metrics e_1, e_2 or e_∞ on $X \times Y$.

A sequence in $X \times Y$ can be written as a sequence of ordered pairs, $((x_n, y_n))$ – that is, for each $n \in \mathbb{N}$, the ordered pair (x_n, y_n) belongs to $X \times Y$, with $x_n \in X$ and $y_n \in Y$. For example, if $X = Y = \mathbb{R}$, then

$$(1, -1),\ (2, 0),\ (3, 1),\ \ldots,\ (n, n-2),\ \ldots$$

is a sequence in $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Definition 3.14 *Component sequences*

Let X and Y be sets and let $((x_n, y_n))$ be a sequence in $X \times Y$. The sequences (x_n) in X and (y_n) in Y are the **component sequences** or **coordinate sequences** of the **product sequence** $((x_n, y_n))$.

The component sequences for the above example are $(1, 2, 3, \ldots, n, \ldots)$ and $(-1, 0, 1, \ldots, n-2, \ldots)$.

Note that the component sequences (x_n) and (y_n) are the images of the original sequence $((x_n, y_n))$ under the projection maps $p_1 \colon X \times Y \to X$ and $p_2 \colon X \times Y \to Y$ given by $p_1((x, y)) = x$ and $p_2((x, y)) = y$.

Since Theorem 3.11 tells us that the projection map p_i is (e, d_i)-continuous for $i = 1, 2$, we can immediately conclude that if $((x_n, y_n))$ e-converges to $(x, y) \in X \times Y$, then $x_n \xrightarrow{d_1} x$ and $y_n \xrightarrow{d_2} y$ as $n \to \infty$. The following result shows that the converse is also true.

Theorem 3.15

Let (X, d_1) and (Y, d_2) be metric spaces and let e be one of the three product metrics e_1, e_2 and e_∞ for $X \times Y$.

Suppose that $(x, y) \in X \times Y$ and $((x_n, y_n))$ is a sequence in $X \times Y$. Then

$$(x_n, y_n) \xrightarrow{e} (x, y) \quad \text{if, and only if,} \quad x_n \xrightarrow{d_1} x \text{ and } y_n \xrightarrow{d_2} y \text{ as } n \to \infty.$$

Proof We have already noted that if $(x_n, y_n) \xrightarrow{e} (x, y)$, then both $x_n \xrightarrow{d_1} x$ and $y_n \xrightarrow{d_2} y$.

It remains to show the converse. So suppose that $x_n \xrightarrow{d_1} x$ and $y_n \xrightarrow{d_2} y$ as $n \to \infty$. Then both $(d_1(x_n, x))$ and $(d_2(y_n, y))$ are (real) null sequences. But for the product metric e_1,

$$e_1((x_n, y_n), (x, y)) = d_1(x_n, x) + d_2(y_n, y)$$

and so, by the Sum Rule, $(e_1((x_n, y_n), (x, y)))$ is a null sequence, meaning that $((x_n, y_n))$ converges in $(X \times Y, e_1)$ to (x, y).

We now use the equivalence of the three metrics e_1, e_2 and e_∞ and Proposition 3.7 to conclude that $((x_n, y_n))$ converges to (x, y) in $(X \times Y, e)$, where e is any of the three product metrics. ■

By induction, we can prove an analogous result for the product of k metric spaces.

Let (X_1, d_1), (X_2, d_2), …, (X_k, d_k) be metric spaces, and let e be a product metric on the set $X_1 \times X_2 \times \cdots \times X_k$. A sequence $((x_{1,n}, x_{2,n}, \ldots, x_{k,n}))_{n \in \mathbb{N}}$ in the product set $X_1 \times X_2 \times \cdots \times X_k$ converges to $(x_1, x_2, \ldots, x_k)$ with respect to the product metric e if, and only if, each component sequence converges for its own metric: $x_{j,n} \xrightarrow{d_j} x_j$ as $n \to \infty$, for $j = 1, 2, \ldots, k$.

Exercise 3.5

Consider the sequence $((\frac{1}{n}, \frac{1}{n^2}))$ in $\mathbb{R}^2$, and let e be the product metric obtained by taking the product of the Euclidean metric with the discrete metric. Does the sequence converge for e and, if so, to what?

$e((x_1, x_2), (y_1, y_2)) = d^{(1)}(x_1, y_1) + d_0(x_2, y_2)$.

4 Spaces of functions

In this section we continue our study of (real-valued) continuous functions by looking at two different ways of defining convergence for sequences of functions – *pointwise convergence* and *uniform convergence*.

Once we have these notions we can ask whether, if the functions $f_n \colon X \to Y$ are continuous (with respect to some metrics on X and Y), the limit function, if it exists, is necessarily continuous. We show that the notion of pointwise convergence does not guarantee this, but that uniform convergence does.

4.1 Pointwise convergence

We first discuss the simplest way of defining convergence for sequences of real-valued functions – pointwise convergence.

There is no assumption here that A has a metric defined on it.

If A is some fixed set and $f_n\colon A \to \mathbb{R}$ for each $n \in \mathbb{N}$ then, for each $x \in A$, $(f_n(x))$ is a sequence of real numbers. Pointwise convergence involves looking separately at each x in the set A and investigating whether the corresponding sequence of real numbers $(f_n(x))$ converges.

Definition 4.1 *Pointwise convergence*

Let A be a set.

A sequence (f_n) of functions $f_n\colon A \to \mathbb{R}$ **converges pointwise** to the function $f\colon A \to \mathbb{R}$ if, for each $x \in A$, the sequence of real numbers $(f_n(x))$ converges to $f(x)$.

That is, $(|f_n(x) - f(x)|)_{n\in\mathbb{N}}$ is a real null sequence.

If a sequence (f_n) of functions converges pointwise to a function f, then f is called the **pointwise limit** of the sequence (f_n).

We write $f_n \to f$ pointwise as $n \to \infty$, $f_n \to_p f$ as $n \to \infty$ or simply $f_n \to f$ pointwise.

Remarks

1. We have not assumed anything about the continuity of the individual functions f_n. Indeed, since A need not have a metric defined for it, it may not even make sense to talk about the continuity of the functions f_n.
2. If it exists, then we know that the pointwise limit is unique, because for each $x \in A$, the sequence $(f_n(x))$ is a sequence of real numbers, and the convergence of a sequence of real numbers corresponds to convergence in $\mathbb{R}$ with respect to the Euclidean metric. In particular, limits (when they exist) are unique.

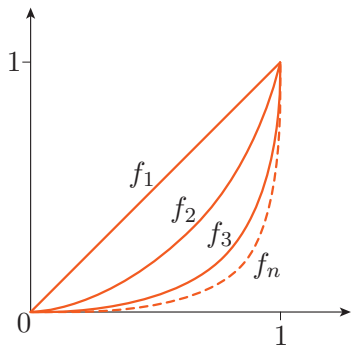


Figure 4.1 The functions f_n

Worked Exercise 4.2

For each $n \in \mathbb{N}$, let $f_n\colon [0,1] \to \mathbb{R}$ be defined by $f_n(x) = x^n$. Find the pointwise limit of the sequence (f_n).

Solution

Since $x^n \to 0$ for each $x \in [0,1)$ but $x^n \to 1$ when $x = 1$, we need to consider the point $x = 1$ separately.

For a fixed $x \in [0,1)$, the sequence of values of the functions at x is $(f_n(x)) = (x^n)$, and this real sequence has limit 0 as $n \to \infty$ (Figure 4.1). Thus, for $x \in [0,1)$, $f_n(x) \to 0$ as $n \to \infty$.

For $x = 1$ and all $n \in \mathbb{N}$, $f_n(1) = 1$, so $f_n(1) \to 1$ as $n \to \infty$.

Hence the pointwise limit of the sequence (f_n) is the function $f\colon [0,1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0, & \text{if } 0 \leq x < 1, \\ 1, & \text{if } x = 1. \end{cases}$$

Notice that each of the functions f_n in Worked Exercise 4.2 is continuous (with respect to the Euclidean metric) – see Figure 4.1 – and yet the limit function is not (see Figure 4.2). This illustrates the fundamental disadvantage of pointwise convergence – it does not generally preserve any 'nice' properties of the functions f_n in the limit. We will see further examples of this shortly, but for now the result is highlighted below.

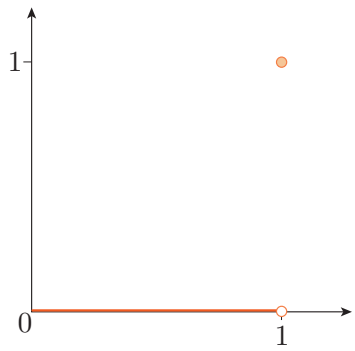


Figure 4.2 The limit function

The pointwise limit of a sequence of continuous functions need not be continuous.

Exercise 4.1

For each $n \in \mathbb{N}$, let $f_n\colon [0,\infty) \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} 1, & \text{if } 0 \leq x \leq \frac{1}{2^n}, \\ -\frac{1}{2^n}, & \text{if } \frac{1}{2^n} < x < \infty. \end{cases}$$

Find the pointwise limit of the sequence of functions (f_n).

Exercise 4.2

For each $n \in \mathbb{N}$, let $f_n\colon [0,1] \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} 4n^2x, & \text{if } 0 \leq x \leq \frac{1}{2n}, \\ -4n^2x + 4n, & \text{if } \frac{1}{2n} < x < \frac{1}{n}, \\ 0, & \text{if } \frac{1}{n} \leq x \leq 1. \end{cases}$$

(a) Sketch the graphs of f_1, f_2 and f_3.

(b) Find the pointwise limit f of the sequence (f_n), and calculate $\int_0^1 f(x)\,dx$.

(c) Find the limit of the real sequence $(\int_0^1 f_n(x)\,dx)_{n\in\mathbb{N}}$.

(d) What conclusion, if any, can be drawn about the relation between integration and pointwise convergence?

4.2 Uniform convergence

We have seen that pointwise convergence of functions need not preserve continuity: a sequence of continuous functions may have a discontinuous pointwise limit. In this subsection, we introduce a type of convergence that does preserve continuity – *uniform convergence*.

You can learn more about Augustin-Louis Cauchy (1789–1857) in the *History Reader*.

Historically, the definition of uniform convergence arose out of a mistake made by Cauchy in his *Lectures on Analysis*, published in 1821. Cauchy was considering an infinite *series* of functions $u_i\colon [0,1] \to \mathbb{R}$, and made the following assertion:

> if the functions u_i are all continuous and if the sequence of partial sums $s_n\colon [0,1] \to \mathbb{R}$ given by
>
> $$s_n(x) = \sum_{i=1}^{n} u_i(x)$$
>
> converges pointwise to a limit function s, then s is also continuous.

The function s_n is continuous, since it is a *finite* sum of continuous functions.

Here $s = \sum_{i=1}^{\infty} u_i$.

In the last subsection, we saw that pointwise convergence does not preserve continuity in general, and it is not difficult to find functions u_n for which the pointwise limit of the partial sums exists and is *not* continuous, meaning that Cauchy's assertion is incorrect. The error was first noticed in 1826 by Abel, who presented examples of sequences of (continuous) partial sums whose pointwise limit is discontinuous.

The *History Reader* says more about Niels Henrik Abel (1802–1829) and his investigation of the convergence of sequences and series.

Let us see where Cauchy's mistake lies. He wrote

$$s(x) = s_n(x) + \sum_{i=n+1}^{\infty} u_i(x),$$

and reasoned as follows. Since s_n converges to s as $n \to \infty$, given any $\varepsilon > 0$ we can choose an $N \in \mathbb{N}$, depending on ε, so that for each $n > N$, the 'tail' of the series is less than ε:

$$\left| \sum_{i=n+1}^{\infty} u_i(x) \right| < \varepsilon.$$

Hence, he wrote, $s(x)$ is a finite sum of continuous functions plus a correction that is less than ε, and so s is continuous.

The issue is that, in general, N depends upon both ε *and* x, and so the whole chain of reasoning just given breaks down. But if the functions u_i are such that we *can* choose an N independently of x, then the result follows. In this case, the same N works *simultaneously* for all x.

We showed in Worked Exercise 4.2 that the limit function $f\colon [0,1] \to \mathbb{R}$ is given by

$$f(x) = \begin{cases} 0, & \text{if } 0 \leq x < 1, \\ 1, & \text{if } x = 1. \end{cases}$$

To see what can go wrong, let us reconsider the convergence of the sequence (f_n) where $f_n\colon [0,1] \to \mathbb{R}$ is given by $f_n(x) = x^n$. We know that the limit function f is discontinuous with respect to the Euclidean metric, and so whatever goes wrong with pointwise convergence should be discernible from this sequence.

Suppose that $0 < \varepsilon < 1$ and $x \in (0,1)$. For what value of n does x^n become less than ε? Solving $x^r = \varepsilon$ for r, we find that $r = \frac{\log \varepsilon}{\log x}$. As anticipated, $r = r(\varepsilon, x)$ depends on both x and ε. Thus, if $N(\varepsilon, x) \in \mathbb{N}$ is chosen so that $N(\varepsilon, x) \geq r(\varepsilon, x)$, then $x^n < \varepsilon$ for each $n > N(\varepsilon, x)$.

The crucial question is whether there is an integer $N(\varepsilon)$, depending on ε but not x, that can serve for a 'uniform rate' of decrease of the x^n over the entire domain $[0,1]$.

$N(\varepsilon)$ needs to be at least as large as all of the numbers $N(\varepsilon, x)$.

The answer is *no*, for as x approaches 1 from below, $N(\varepsilon, x)$ grows without limit; that is,

$$\sup\{N(\varepsilon, x) : x \in [0,1]\} \geq \sup\{\tfrac{\log \varepsilon}{\log x} : x \in (0,1)\} = \infty.$$

The hope is that if we require this *not* to happen – that is, if the set $\{N(\varepsilon, x) : x \in [0,1]\}$ is bounded above for each $\varepsilon > 0$ – then we obtain a 'uniform control', and the limit function of a sequence of continuous functions *is* continuous. This is the idea underlying the notion of *uniform convergence*.

Before we define uniform convergence, we need to extend our definition of a bounded function to situations when the domain may not be a subset of $\mathbb{R}$.

Definition 4.3 *Bounded function*

Let A be a set. A function $f\colon A \to \mathbb{R}$ is **bounded** on A if there is an $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in A$.

This extends the definition of boundedness of a function, which you met in Definition 2.14 of Chapter 13.

We can now give the definition of uniform convergence.

Definition 4.4 *Uniform convergence*

Let A be a set. A sequence (f_n) of functions $f_n\colon A \to \mathbb{R}$ **converges uniformly** on A to the function $f\colon A \to \mathbb{R}$ if

(a) the function $f_n - f$ is bounded for each $n \in \mathbb{N}$, and

(b) the sequence (M_n) defined by

$$M_n = \sup\{|f_n(x) - f(x)| : x \in A\}$$

is a (real) null sequence.

Such a function f is called the **uniform limit** of the sequence (f_n).

We write either $f_n \to f$ uniformly as $n \to \infty$, $f_n \rightrightarrows f$ as $n \to \infty$ or simply $f_n \to f$ uniformly.

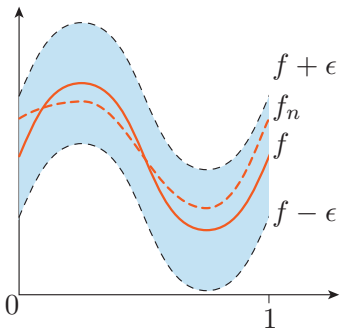


Figure 4.3 Uniform convergence

Remarks

1. The requirement that $f_n - f$ is bounded is imposed to ensure that M_n is a well-defined real number for each $n \in \mathbb{N}$.
2. As with the definition of pointwise convergence, we make no assumption as to whether there is a metric for the domain of the functions.
3. We must always specify on which domain the sequence converges uniformly. A sequence might converge uniformly on A but not on some larger set $B \supset A$. This is illustrated in Worked Exercise 4.6 and Exercise 4.4 (below).
4. There is a useful visual aid to understanding the meaning of uniform convergence when $A = [0, 1]$. Given $\varepsilon > 0$, we draw a 'sleeve' of total width 2ε symmetrically around the graph of f (Figure 4.3). If, for each $\varepsilon > 0$, there is an $N(\varepsilon) \in \mathbb{N}$ such that the graph of each function f_n for $n > N(\varepsilon)$ fits inside this sleeve over the whole of A, then the convergence is uniform. If there is at least one ε for which we cannot find such a sleeve, then the convergence is not uniform.

Before we look at some examples of uniform convergence, we note a useful lemma that relates uniform and pointwise convergence of sequences of functions: it states that uniform convergence implies pointwise convergence.

Lemma 4.5

Let A be a set and let (f_n) be a sequence of functions $f_n \colon A \to \mathbb{R}$.

If the sequence (f_n) converges uniformly on A to the function $f \colon A \to \mathbb{R}$, then (f_n) also converges pointwise to f.

Proof Suppose that (f_n) converges uniformly to f. We must show that for each $x \in A$, the real sequence $(f_n(x))$ converges to $f(x)$.

For each $n \in \mathbb{N}$, let

$$M_n = \sup\{|f_n(x) - f(x)| : x \in A\}$$

and note that, since f_n converges uniformly to f, $M_n < \infty$ for each n and (M_n) is a real null sequence.

But for each $n \in \mathbb{N}$ and each $x \in A$,

$$0 \leq |f_n(x) - f(x)| \leq M_n.$$

See Lemma 1.15 of Chapter 13.

Hence, by the Squeeze Rule for sequences, $(f_n(x) - f(x))_{n \in \mathbb{N}}$ is a null sequence, and so $f_n(x) \to f(x)$ as $n \to \infty$. That is, f_n converges to f pointwise. ■

This lemma tells us that if a sequence of functions converges uniformly to a function, then it also converges pointwise to the same function. Thus, in order to identify a possible uniform limit of a sequence of functions, it is enough to find the (unique) pointwise limit of the sequence. Having found this candidate for the uniform limit, we then have to verify whether or not it *is* the uniform limit. The following worked exercise illustrates the method.

Worked Exercise 4.6

For $n \in \mathbb{N}$, let $f_n : [0,1] \to \mathbb{R}$ be given by

$$f_n(x) = \tfrac{x}{n}.$$

Show that the sequence (f_n) converges uniformly on $[0,1]$, and find its uniform limit.

Solution

We begin by identifying the pointwise limit of the sequence. For each $x \in [0,1]$, $\frac{x}{n} \to 0$ as $n \to \infty$, and so the pointwise limit of the sequence (f_n) is the zero function $f : [0,1] \to \mathbb{R}$ given by $f(x) = 0$ for each $x \in [0,1]$.

Thus, by Lemma 4.5, our only candidate for the uniform limit of the sequence (f_n) is the function f. To show that f is the uniform limit of (f_n), we must show that the sequence (M_n), where M_n is defined by

$$M_n = \sup\{|f_n(x) - f(x)| : x \in [0,1]\},$$

is a null sequence.

To prove this, we observe that for each $x \in [0,1]$ and each $n \in \mathbb{N}$,

$$0 \leq |f_n(x) - f(x)| = |\tfrac{x}{n} - 0| = |\tfrac{x}{n}| \leq \tfrac{1}{n}.$$

Hence for each $n \in \mathbb{N}$,

$$0 \leq M_n = \sup\{|\tfrac{x}{n} - 0| : x \in [0,1]\} = \tfrac{1}{n}.$$

Since $(\frac{1}{n})$ is a basic null sequence, it follows that (M_n) is a null sequence. Hence $f_n \to f$ uniformly as $n \to \infty$.

Exercise 4.3

For $n \in \mathbb{N}$, let $f_n : [0,1] \to \mathbb{R}$ be given by

$$f_n(x) = 1 - (\tfrac{x}{n})^2.$$

Show that the sequence of functions (f_n) converges uniformly on $[0,1]$, and find its uniform limit.

Exercise 4.4

Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by

$$f_n(x) = \frac{x}{n}.$$

Show that the sequence of functions (f_n) does not converge uniformly on $\mathbb{R}$.

We have seen in Worked Exercise 4.6 and Exercise 4.4 that the functions given by $f_n(x) = \frac{x}{n}$ converge uniformly on $[0, 1]$, but do *not* converge uniformly on the whole of $\mathbb{R}$. Exercise 4.4 also illustrates another point.

Even when a sequence converges pointwise to a *continuous* function, the convergence need not be uniform.

Uniform convergence and continuity

We now show that uniform convergence preserves continuity – that is, *the uniform limit of a sequence of continuous functions is also continuous.*

Recall that in order to consider continuity, we need a metric on the domain as well as a metric on the codomain $\mathbb{R}$. We choose to use the usual Euclidean metric for the codomain, $d^{(1)}(x, y) = |x - y|$.

Theorem 4.7

Let (X, d) be a metric space and fix $a \in X$.

For each $n \in \mathbb{N}$, suppose that $f_n : X \to \mathbb{R}$ is $(d, d^{(1)})$-continuous at a.

If the sequence of functions (f_n) converges uniformly on X to a function $f : X \to \mathbb{R}$, then f is $(d, d^{(1)})$-continuous at a.

Proof The proof of this result takes some care.

Let (X, d) be a metric space and fix $a \in X$. For each $n \in \mathbb{N}$, suppose that $f_n : X \to \mathbb{R}$ is $(d, d^{(1)})$-continuous at a and that the sequence of functions (f_n) converges uniformly to the function $f : X \to \mathbb{R}$.

Our goal is to show that f is $(d, d^{(1)})$-continuous at a; that is, if (x_m) is a sequence in X that d-converges to a, then $f(x_m) \to f(a)$ as $m \to \infty$.

Notice that the sequence is indexed by a different parameter from that used for the sequence of functions.

To do this, it is enough to show that for each $\varepsilon > 0$, there is $N \in \mathbb{N}$ for which $|f(x_m) - f(a)| < \varepsilon$ whenever $m > N$.

Fix $\varepsilon > 0$.

Since f_n converges uniformly to f as $n \to \infty$, we know that the sequence (M_n) defined by

$$M_n = \sup\{|f_n(x) - f(x)| : x \in X\}$$

is a real null sequence. Since this is the only information we have about the relationship between the functions f_n and f, it should not be a surprise that we use it in the proof.

Let us start by observing that for any term x_m of the sequence, we can use the Triangle Inequality for the modulus function to estimate that for each $m \in \mathbb{N}$ and $n \in \mathbb{N}$,

$$\begin{aligned} |f(x_m)-f(a)| &= |f(x_m) - f_n(x_m) + f_n(x_m) - f_n(a) + f_n(a) - f(a)| \\ &= |(f(x_m) - f_n(x_m)) + (f_n(x_m) - f_n(a)) + (f_n(a) - f(a))| \\ &\leq |f(x_m) - f_n(x_m)| + |f_n(x_m) - f_n(a)| + |f_n(a) - f(a)| \\ &\leq M_n + |f_n(x_m) - f_n(a)| + M_n \\ &= 2M_n + |f_n(x_m) - f_n(a)|. \end{aligned}$$

This splitting of an expression by adding and subtracting identical terms is a very useful technique in analysis.

However, we know that (M_n) is a real null sequence (of non-negative terms) and so, in particular, we can find $n \in \mathbb{N}$ for which $0 \leq M_n < \frac{\varepsilon}{3}$. But we also know that for this choice of n the function f_n is $(d, d^{(1)})$-continuous at a and so, since $x_m \xrightarrow{d} a$, there is $N \in \mathbb{N}$ for which $d^{(1)}(f_n(x_m), f_n(a)) < \frac{\varepsilon}{3}$ whenever $m > N$. We conclude that for this choice of N, whenever $m > N$,

$$|f(x_m) - f(a)| < 2 \times \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} = \varepsilon,$$

as required.

Hence f is $(d, d^{(1)})$-continuous at a. ■

Theorems such as this one are very useful: they show that under certain circumstances, the limit of a sequence can have some of the properties of the terms of the sequence.

Exercise 4.5

For each $n \in \mathbb{N}$, let $f_n \colon [0, \infty) \to \mathbb{R}$ be given by

$$f_n(x) = \begin{cases} 1, & \text{if } x = 0, \\ \dfrac{\sin(nx)}{nx}, & \text{if } x > 0. \end{cases}$$

(a) Does this sequence converge pointwise?

(b) Does it converge uniformly? If it does converge, find its limit function.

Hint: $\lim_{y \to 0} \frac{\sin y}{y} = 1$.

5 The max metric on $C[0, 1]$

In this final section of the chapter, we introduce a metric for a particular set of functions that is based on ideas related to uniform convergence. In this example, the functions in question are the continuous functions from the closed interval $[0, 1]$ to $\mathbb{R}$, and the metric we define is known as the *max metric*.

Notation

We let $C[0, 1]$ denote the set consisting of all functions $f\colon [0, 1] \to \mathbb{R}$ that are continuous on $[0, 1]$ with respect to the Euclidean metrics on $[0, 1]$ and $\mathbb{R}$.

Thus for each $a \in [0, 1]$ and each sequence (x_k) in $[0, 1]$, if $|x_k - a| \to 0$ as $k \to \infty$, then $|f(x_k) - f(a)| \to 0$ as $k \to \infty$.

Remarks

1. The points of $C[0, 1]$ are continuous functions on $[0, 1]$ such as $f\colon [0, 1] \to \mathbb{R}$ given by $f(x) = 3x$ and $g\colon [0, 1] \to \mathbb{R}$ given by $g(x) = \sin x$, for $x \in [0, 1]$.
2. We define the set $C[a, b]$ of all continuous functions from the closed interval $[a, b]$ to $\mathbb{R}$ in a similar way.

These properties can be deduced from the combination rules for continuous functions from $\mathbb{R}$ to $\mathbb{R}$, together with the Restriction Rule (Proposition 2.8) from Chapter 13.

In discussing the max metric, we are going to need the following properties of functions in $C[0, 1]$, all of which can be easily deduced from corresponding results given in Chapter 13.

Properties of functions in $C[0, 1]$

Let f and g be real-valued continuous functions with domain $[0, 1]$, i.e. $f, g \in C[0, 1]$, and let $\lambda \in \mathbb{R}$. Then the following real-valued functions are all continuous on $[0, 1]$, i.e. they are all members of $C[0, 1]$:

Sum Rule	$f + g\colon [0, 1] \to \mathbb{R}$, defined by $(f + g)(x) = f(x) + g(x)$
Multiple Rule	$\lambda f\colon [0, 1] \to \mathbb{R}$, defined by $(\lambda f)(x) = \lambda \times f(x)$
Product Rule	$fg\colon [0, 1] \to \mathbb{R}$, defined by $(fg)(x) = f(x)g(x)$
Modulus Rule	$\lvert f\rvert\colon [0, 1] \to \mathbb{R}$, defined by $\lvert f\rvert(x) = \lvert f(x)\rvert$.

A *real-valued function* is one with codomain $\mathbb{R}$.

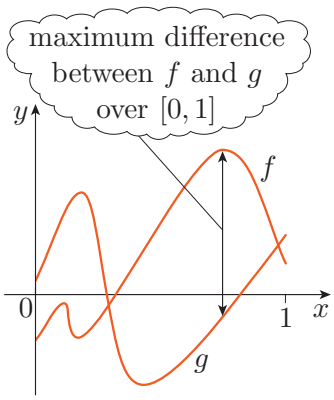


Figure 5.1 The max metric measures the maximum difference between two functions

The metric we define on $C[0, 1]$ is based on the idea of measuring the distance between two functions f and g in $C[0, 1]$ in terms of their maximum difference over the interval $[0, 1]$. This is given by

$$\max\{|g(x) - f(x)| : x \in [0, 1]\}$$

and is illustrated in Figure 5.1. It makes sense to talk of this quantity, because the map $x \mapsto |g(x) - f(x)|$ defines a continuous function on $[0, 1]$ and hence, by the Extreme Value Theorem (Theorem 2.16 of Chapter 13), this maximum exists. We can thus make the following definition.

Definition 5.1 *Max metric on $C[0,1]$*

The **max metric** on $C[0,1]$ is the function $d_{\max}\colon C[0,1] \times C[0,1] \to \mathbb{R}$ defined by

$$d_{\max}(f,g) = \max\{|g(x) - f(x)| : x \in [0,1]\}.$$

We show that $d_{\max}$ is a metric shortly.

Exercise 5.1

Find $d_{\max}(f,g)$ when $f, g\colon [0,1] \to \mathbb{R}$ are given by:

(a) $f(x) = 1$, $g(x) = x$ for $0 \leq x \leq 1$

(b) $f(x) = \frac{1}{2}$, $g(x) = \sin 2\pi x$ for $0 \leq x \leq 1$

(c) $f(x) = x$, $g(x) = x^2$ for $0 \leq x \leq 1$.

Hint: in answering part (c), you might find it helpful to write $h = g - f$ and to recall that the maximum and minimum of a function h can be found by considering $h(x)$ at the points x where $h'(x) = 0$ and at the endpoints $x = 0$, $x = 1$. You may also find it helpful to sketch the graph of h.

We now show that $d_{\max}$ is indeed a metric on $C[0,1]$.

Proposition 5.2

$(C[0,1], d_{\max})$ is a metric space.

Proof We show that $d_{\max}$ satisfies conditions (M1)–(M3).

(M1) Since $|g(x) - f(x)| \geq 0$ for each $x \in [0,1]$, its maximum is also non-negative. So $d_{\max}(f,g) \geq 0$ for any f and g in $C[0,1]$.

For each $f \in C[0,1]$ and each $x \in [0,1]$, we have $|f(x) - f(x)| = 0$; therefore $d_{\max}(f,f) = 0$ whenever $f \in C[0,1]$.

If f and g in $C[0,1]$ are such that $d_{\max}(f,g) = 0$ then, for each $x \in [0,1]$, we have

$$0 \leq |g(x) - f(x)| \leq d_{\max}(f,g) = 0;$$

hence $|g(x) - f(x)| = 0$ for each $x \in [0,1]$. Thus, $f(x) = g(x)$ for each $x \in [0,1]$, which implies that $f = g$.

So $d_{\max}$ satisfies (M1).

(M2) For any $f, g \in C[0,1]$ and each $x \in [0,1]$, property (M2) for the modulus function tells us that

$$|f(x) - g(x)| = |g(x) - f(x)|,$$

and so

$$d_{\max}(g, f) = \max\{|f(x) - g(x)| : x \in [0, 1]\}$$
$$= \max\{|g(x) - f(x)| : x \in [0, 1]\} = d_{\max}(f, g).$$

Thus $d_{\max}$ satisfies (M2).

(M3) Let $f, g, h \in C[0, 1]$ and $x \in [0, 1]$. Since the modulus function has property (M3), we have

$$|h(x) - f(x)| \leq |h(x) - g(x)| + |g(x) - f(x)|. \quad (1)$$

But for each $x \in [0, 1]$,

$$|h(x) - g(x)| \leq \max\{|h(x) - g(x)| : x \in [0, 1]\}$$
$$= d_{\max}(g, h), \quad (2)$$

and, similarly,

$$|g(x) - f(x)| \leq d_{\max}(f, g). \quad (3)$$

Thus for each $x \in [0, 1]$, by combining equations (1), (2) and (3), we see that

$$|h(x) - f(x)| \leq d_{\max}(f, g) + d_{\max}(g, h).$$

Since this is true for each $|h(x) - f(x)|$ with $x \in [0, 1]$, it is true for the maximum value, and so

$$d_{\max}(f, h) \leq d_{\max}(f, g) + d_{\max}(g, h).$$

Thus $d_{\max}$ satisfies (M3).

Since $d_{\max}$ satisfies conditions (M1)–(M3), it is a metric on $C[0, 1]$. Hence $(C[0, 1], d_{\max})$ is a metric space. ■

In a similar way, for any interval $[a, b]$, we can define a metric $d_{\max}$ on $C[a, b]$ by

$$d_{\max}(f, g) = \max\{|g(x) - f(x)| : x \in [a, b]\}.$$

Exercise 5.2

For $f, g \in C[0, 1]$, define $d_{\min}(f, g)$ by

$$d_{\min}(f, g) = \min\{|g(x) - f(x)| : x \in [0, 1]\}.$$

Determine whether $d_{\min}$ is a metric on $C[0, 1]$.

The next worked exercise gives another important example of a metric for $C[0, 1]$.

Worked Exercise 5.3 *Integral metric*

This worked exercise investigates a metric on the set $C[0,1]$ that is different from $d_{\max}$.

Define $d\colon C[0,1] \times C[0,1] \to \mathbb{R}$ by

$$d(f,g) = \int_0^1 |g(x) - f(x)|\,dx.$$

(a) Find $d(f,g)$ when $f, g\colon [0,1] \to \mathbb{R}$ are given by

(i) $f(x) = x$, $g(x) = 0$; (ii) $f(x) = x$, $g(x) = x^2$;
(iii) $f(x) = \sin 2\pi x$, $g(x) = 0$.

(b) Prove that d defines a metric on $C[0,1]$.

Solution

(a) (i) Here $|g(x) - f(x)| = |0 - x| = x$ when $x \in [0,1]$. Hence

$$d(f,g) = \int_0^1 x\,dx = \left[\tfrac{1}{2}x^2\right]_0^1 = \tfrac{1}{2}.$$

(ii) Here $|g(x) - f(x)| = |x^2 - x| = x - x^2$ when $x \in [0,1]$. Hence

$$d(f,g) = \int_0^1 (x - x^2)\,dx = \left[\tfrac{1}{2}x^2 - \tfrac{1}{3}x^3\right]_0^1 = \tfrac{1}{6}.$$

(iii) Here $|g(x) - f(x)| = |0 - \sin 2\pi x| = |\sin 2\pi x|$ when $x \in [0,1]$.

Now $|\sin 2\pi x| = \sin 2\pi x$ when $x \in [0, \frac{1}{2}]$, and
$|\sin 2\pi x| = -\sin 2\pi x$ when $x \in (\frac{1}{2}, 1]$. Hence

$$\begin{aligned} d(f,g) &= \int_0^{\frac{1}{2}} \sin 2\pi x\,dx - \int_{\frac{1}{2}}^1 \sin 2\pi x\,dx \\ &= \left[\frac{-1}{2\pi}\cos 2\pi x\right]_0^{\frac{1}{2}} - \left[\frac{-1}{2\pi}\cos 2\pi x\right]_{\frac{1}{2}}^1 = \frac{2}{\pi}. \end{aligned}$$

(b) To show that d is a metric on $C[0,1]$, we must verify that d satisfies (M1)–(M3).

(M1) Let $f, g \in C[0,1]$. Since $|g(x) - f(x)| \geq 0$ whenever $x \in [0,1]$, the monotonicity property of integration implies that

$$d(f,g) = \int_0^1 |g(x) - f(x)|\,dx \geq \int_0^1 0\,dx = 0.$$

Basic properties of integrals are summarised in the *Handbook*.

Direct calculation gives

$$d(f,f) = \int_0^1 |f(x) - f(x)|\,dx = \int_0^1 0\,dx = 0.$$

If $f, g \in C[0,1]$ satisfy $d(f,g) = 0$, then

$$\int_0^1 |g(x) - f(x)|\,dx = 0,$$

See the *Handbook*.

and so, since $x \mapsto |g(x) - f(x)|$ is a non-negative continuous function on $[0,1]$, it follows that $|g(x) - f(x)| = 0$ for each $x \in [0,1]$. Hence $f(x) = g(x)$ for each $x \in [0,1]$, and so $f = g$.

Thus d satisfies (M1).

(M2) Let $f, g \in C[0,1]$. Then

$$\begin{aligned} d(f,g) &= \int_0^1 |g(x) - f(x)|\,dx \\ &= \int_0^1 |f(x) - g(x)|\,dx = d(g,f), \end{aligned}$$

and so d satisfies (M2).

(M3) Let $f, g, h \in C[0,1]$. Using the Triangle Inequality for the modulus function together with the monotonicity property of integration, we have

$$\begin{aligned} d(f,h) &= \int_0^1 |h(x) - f(x)|\,dx \\ &= \int_0^1 |(h(x) - g(x)) + (g(x) - f(x))|\,dx \\ &\leq \int_0^1 (|h(x) - g(x)| + |g(x) - f(x)|)\,dx \\ &= \int_0^1 |h(x) - g(x)|\,dx + \int_0^1 |g(x) - f(x)|\,dx \\ &= d(f,g) + d(g,h). \end{aligned}$$

So d satisfies (M3).

We conclude that d is a metric on $C[0,1]$.

Open and closed balls in the max metric

Given a function $f \in C[0,1]$ and $r > 0$, the open ball $B_{d_{\max}}(f,r)$ is determined by finding the set of all functions $g \in C[0,1]$ satisfying the inequality $d_{\max}(f,g) < r$ – that is,

$$d_{\max}(f,g) = \max\{|g(x) - f(x)| : x \in [0,1]\} < r.$$

For each $x \in [0,1]$, if $|g(x) - f(x)| < r$ then $g(x)$ must lie strictly within a distance r of $f(x)$. And so, if $g \in B_{d_{\max}}(f,r)$ then for each $x \in [0,1]$, the value $g(x)$ must lie within a distance r of $f(x)$.

Thus, over the whole of $[0,1]$, $g(x)$ must lie within a band following the shape of f and extending r above and below it at each point, as Figure 5.2 illustrates. It follows that the open ball $B_{d_{\max}}(f,r)$ consists of all those $g \in C[0,1]$ whose graphs lie within this band.

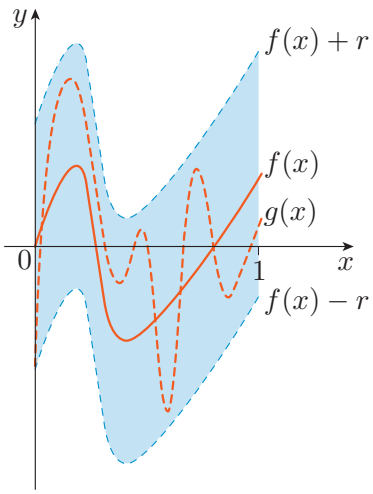


Figure 5.2 The open ball $B_{d_{\max}}(f,r)$ consists of all continuous functions whose graphs lie within the open band illustrated

The closed ball $B_{d_{\max}}[f,r]$ is obtained in a similar way except that the band includes the points exactly distance r above and below the graph of f.

Continuity and the max metric on $C[0,1]$

We now look at some functions defined on $C[0,1]$ and investigate whether they are continuous. Such functions will be a mapping from $C[0,1]$ to another metric space (Y,d). If $(Y,d)=(C[0,1],d_{\max})$, then such a mapping will map continuous functions to continuous functions. A simple example of such a mapping from $C[0,1]$ to $C[0,1]$ is given as follows.

We use the word 'map' to help us distinguish between functions *in* $C[0,1]$ and functions defined *on* $C[0,1]$.

Let us add a fixed function h from $C[0,1]$ to a function $f\in C[0,1]$; then the new function, given by $(f+h)(x)=f(x)+h(x)$ for each $x\in[0,1]$, is also continuous (being the sum of two continuous functions).

Thus we may define a map $F_h\colon C[0,1]\to C[0,1]$ by $f\mapsto(f+h)$ for $f\in C[0,1]$.

Note the use of upper case F to define this map: this helps us to distinguish it from members of $C[0,1]$.

So for example, if $h\in C[0,1]$ is given by $h(x)=x^2$, then we define the map F_h by $f\mapsto f+h$ for $f\in C[0,1]$. Thus, for each $f\in C[0,1]$, the function $F_h(f)$ is a function on $[0,1]$ whose value at a particular point $x\in[0,1]$ is given by

$$F_h(f)(x)=f(x)+x^2.$$

This is shown in Figure 5.3. Since f and h are in $C[0,1]$, $f+h$ is also in $C[0,1]$ and so $F_h\colon C[0,1]\to C[0,1]$.

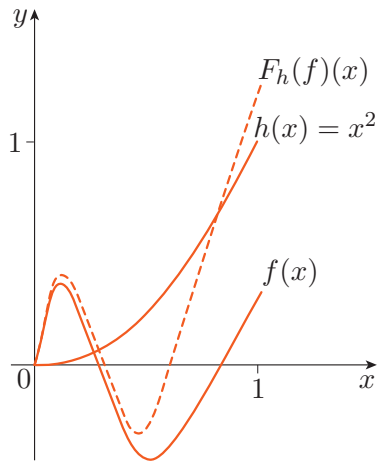


Figure 5.3 Graphs of $f(x)$, $h(x)$ and $F_h(f)(x)$

Worked Exercise 5.4

Show that for each $h\in C[0,1]$, $F_h\colon C[0,1]\to C[0,1]$ is $(d_{\max},d_{\max})$-continuous on $C[0,1]$.

Solution

Fix $h\in C[0,1]$. We show that F_h is $(d_{\max},d_{\max})$-Lipschitz and so, by Theorem 3.2, is $(d_{\max},d_{\max})$-continuous on $C[0,1]$.

Let $f,g\in C[0,1]$ – we must find an upper bound for $d_{\max}(F_h(f),F_h(g))$ in terms of $d_{\max}(f,g)$.

By definition,

$$d_{\max}(F_h(f),F_h(g))=\max\{|F_h(f)(x)-F_h(g)(x)|:x\in[0,1]\}.$$

We consider the value of $|F_h(f)(x)-F_h(g)(x)|$ for a *particular* $x\in[0,1]$. We find that

$$\begin{aligned}|F_h(f)(x)-F_h(g)(x)|&=|(f(x)+h(x))-(g(x)+h(x))|\\&=|f(x)-g(x)|.\end{aligned}$$

Hence, for each $x\in[0,1]$,

$$|F_h(f)(x)-F_h(g)(x)|=|f(x)-g(x)|$$

and so

$$\max\{|F_h(f)(x) - F_h(g)(x)| : x \in [0,1]\} = \max\{|f(x) - g(x)| : x \in [0,1]\}.$$

That is,

$$d_{\max}(F_h(g), F_h(f)) = d_{\max}(g, f).$$

Thus F_h is $(d_{\max}, d_{\max})$-Lipschitz (with Lipschitz constant 1) and the result follows.

We have seen that adding a fixed continuous function to a function in $C[0,1]$ produces a continuous mapping from $(C[0,1], d_{\max})$ to itself. Similarly, multiplying by a fixed continuous function also produces a continuous mapping from $(C[0,1], d_{\max})$ to itself: we choose a function $h \in C[0,1]$ and define the *product* function by $G_h(f) = hf$, for $f \in C[0,1]$. Since f and h are in $C[0,1]$, the product hf is also in $C[0,1]$ and so $G_h : C[0,1] \to C[0,1]$.

Thus $G_h(f)(x) = h(x) \times f(x)$.

Worked Exercise 5.5

Show that for each $h \in C[0,1]$, G_h is $(d_{\max}, d_{\max})$-continuous on $C[0,1]$.

Solution

The proof is very similar to that of Worked Exercise 5.4.

Fix $h \in C[0,1]$.

We show that G_h is $(d_{\max}, d_{\max})$-Lipschitz and so, by Theorem 3.2, is $(d_{\max}, d_{\max})$-continuous on $C[0,1]$. Let $f, g \in C[0,1]$ – we must find an upper bound for $d_{\max}(G_h(f), G_h(g))$ in terms of $d_{\max}(f, g)$.

As in Worked Exercise 5.4, we know that

$$d_{\max}(G_h(f), G_h(g)) = \max\{|G_h(f)(x) - G_h(g)(x)| : x \in [0,1]\}.$$

We consider the value of $|G_h(f)(x) - G_h(g)(x)|$ for a *particular* $x \in [0,1]$. We find that

$$\begin{aligned} |G_h(f)(x) - G_h(g)(x)| &= |h(x)f(x) - h(x)g(x)| \\ &= |h(x)(f(x) - g(x))| \\ &= |h(x)|\,|f(x) - g(x)| \leq |h(x)|\, d_{\max}(g, f). \end{aligned}$$

So we must find an upper bound for $|h(x)|$ in which x does not appear. But $h \in C[0,1]$, and so $|h| \in C[0,1]$. Hence, by the Boundedness Theorem, there is a real number $M \geq 0$ for which

$$|h(x)| \leq M \text{ for each } x \in [0,1].$$

This is Theorem 2.15 in Chapter 13.

It follows that for each $x \in [0,1]$,

$$|h(x)f(x) - h(x)g(x)| \leq M d_{\max}(f,g)$$

and so

$$d_{\max}(G_h(f), G_h(g)) \leq M d_{\max}(f,g).$$

Thus G_h is $(d_{\max}, d_{\max})$-Lipschitz (with Lipschitz constant M, the bound on h) and the result follows.

We now look at a few examples where our mappings go from $C[0,1]$ to a different metric space, $(\mathbb{R}, d^{(1)})$. We again make frequent use of Theorem 3.2 and show that our mappings are Lipschitz in order to show that they are continuous.

Worked Exercise 5.6

Show that $I\colon C[0,1] \to \mathbb{R}$ defined by $I(f) = \int_0^1 f(x)\,dx$ is $(d_{\max}, d^{(1)})$-continuous.

Solution

We show that I is a $(d_{\max}, d^{(1)})$-Lipschitz function and so, by Theorem 3.2, it is $(d_{\max}, d^{(1)})$-continuous.

Let $f, g \in C[0,1]$ be given. We aim to find an upper bound for $d^{(1)}(I(f), I(g))$ in terms of $d_{\max}(f,g)$. Now

$$\begin{aligned} d^{(1)}(I(f), I(g)) &= \left| \int_0^1 g(x)\,dx - \int_0^1 f(x)\,dx \right|, \text{ using definition of } I \\ &= \left| \int_0^1 (g(x) - f(x))\,dx \right|. \end{aligned} \tag{4}$$

The definition of $d_{\max}$ implies that for each $x \in [0,1]$,

$$|g(x) - f(x)| \leq d_{\max}(f,g),$$

which means that for each $x \in [0,1]$,

$$-d_{\max}(f,g) \leq g(x) - f(x) \leq d_{\max}(f,g).$$

Integrating the terms in this double inequality between 0 and 1, we obtain

$$\int_0^1 -d_{\max}(f,g)\,dx \leq \int_0^1 (g(x) - f(x))\,dx \leq \int_0^1 d_{\max}(f,g)\,dx;$$

that is,

$$-d_{\max}(f,g) \leq \int_0^1 (g(x) - f(x))\,dx \leq d_{\max}(f,g),$$

Note that $\int_0^1 d_{\max}(f,g)\,dx$ is just the integral of a constant between 0 and 1 and so takes the value of the constant.

which is equivalent to

$$\left| \int_0^1 (g(x) - f(x))\, dx \right| \leq d_{\max}(f, g).$$

This, together with equation (4), gives

$$d^{(1)}(I(f), I(g)) \leq d_{\max}(f, g).$$

Hence I is $(d_{\max}, d^{(1)})$-Lipschitz (with Lipschitz constant 1), and so by Theorem 3.2 I is $(d_{\max}, d^{(1)})$-continuous on $C[0, 1]$.

Exercise 5.3

Let $h \in C[0, 1]$ be fixed. Define $I_h \colon C[0, 1] \to \mathbb{R}$ by

$$I_h(f) = \int_0^1 h(x) f(x)\, dx.$$

Show that I_h is $(d_{\max}, d^{(1)})$-continuous on $C[0, 1]$.

Hint: you may find Proposition 4.4 of Chapter 14, the Composition Rule, useful.

Worked Exercise 5.7

For $f \in C[0, 1]$, define a function $F(f) \colon [0, 1] \to \mathbb{R}$ by

$$F(f)(t) = \int_0^t f(x)\, dx.$$

(a) Show that $F(f) \in C[0, 1]$.

(b) Show that the function $F \colon C[0, 1] \to C[0, 1]$ is $(d_{\max}, d_{\max})$-continuous.

Hint: show that $|F(f)(t) - F(f)(s)| \leq |t - s|\, d_{\max}(f, 0)$ for each $s, t \in [0, 1]$.

Solution

(a) Let $f \in C[0, 1]$ and $s, t \in [0, 1]$ with $s \leq t$. Then

$$\begin{aligned} |F(f)(t) - F(f)(s)| &= \left| \int_0^t f(x)\, dx - \int_0^s f(x)\, dx \right| \\ &= \left| \int_s^t f(x)\, dx \right| \leq \int_s^t |f(x)|\, dx \\ &\leq \int_s^t d_{\max}(f, 0)\, dx \\ &= (t - s) d_{\max}(f, 0) = |t - s|\, d_{\max}(f, 0). \end{aligned}$$

Hence $F(f)$ is a $(d^{(1)}, d^{(1)})$-Lipschitz function on $[0,1]$ (with Lipschitz constant $d_{\max}(f,0)$), and so $F(f)$ is $(d^{(1)}, d^{(1)})$-continuous by Theorem 3.2; that is, $F(f) \in C[0,1]$.

(b) In order to show that F is $(d_{\max}, d_{\max})$-continuous, let $f, g \in C[0,1]$, and observe that for each $t \in [0,1]$,

$$\begin{aligned} |F(f)(t) - F(g)(t)| &= \left| \int_0^t (f(x) - g(x))\, dx \right| \\ &\leq \int_0^t |f(x) - g(x)|\, dx \\ &\leq \int_0^1 d_{\max}(f,g)\, dx = d_{\max}(f,g). \end{aligned}$$

Hence

$$d_{\max}(F(f), F(g)) \leq d_{\max}(f,g);$$

thus F is a $(d_{\max}, d_{\max})$-Lipschitz function and so, in particular, is a $(d_{\max}, d_{\max})$-continuous function.

5.1 Sequences in $C[0,1]$

We finish this chapter by looking at the relationship of $d_{\max}$-convergence of sequences in $C[0,1]$ to pointwise and uniform convergence.

Recall that $f \in C[0,1]$ if $f\colon [0,1] \to \mathbb{R}$ and f is continuous on $[0,1]$ for the Euclidean metric. The distance $d_{\max}$ between two functions f and g in $C[0,1]$ is defined by

$$d_{\max}(f,g) = \max\{|g(x) - f(x)| : x \in [0,1]\}$$

(we saw in Proposition 5.2 that this defines a metric on $C[0,1]$).

Definition 3.2 of Chapter 14 tells us that a sequence of functions (f_n) in $C[0,1]$ converges to f in $C[0,1]$ for $d_{\max}$ if, and only if, $d_{\max}(f_n, f) \to 0$ as $n \to \infty$; that is if, and only if,

$$\max\{|f(x) - f_n(x)| : x \in [0,1]\} \to 0 \text{ as } n \to \infty.$$

The resemblance between this and the definition of uniform convergence means that the following result is not a great surprise.

Theorem 5.8 *Uniform Convergence Theorem*

Let (f_n) be a sequence of functions in $C[0,1]$.

If (f_n) converges uniformly on $[0,1]$ to the function $f\colon [0,1] \to \mathbb{R}$, then f is continuous on $[0,1]$ (and so is in $C[0,1]$) and $d_{\max}(f_n, f) \to 0$ as $n \to \infty$.

Conversely, if $f \in C[0,1]$ and $d_{\max}(f_n, f) \to 0$ as $n \to \infty$, then (f_n) converges uniformly to f on $[0,1]$.

Proof Let (f_n) be a sequence in $C[0,1]$ that converges uniformly to the function $f\colon [0,1] \to \mathbb{R}$. By Theorem 4.7 of this chapter, f is continuous on $[0,1]$. It remains to show that $(d_{\max}(f_n, f))_{n\in\mathbb{N}}$ is a null sequence.

Since $f_n \to f$ uniformly on $[0,1]$, we know (M_n) is a real null sequence when M_n is given by

$$M_n = \sup\{|f_n(x) - f(x)| : x \in [0,1]\}.$$

In addition, since $[0,1]$ is a closed bounded interval, the Extreme Value Theorem (Theorem 2.16 of Chapter 13) tells us that $f_n - f$ attains its bounds for each $n \in \mathbb{N}$. Thus

$$M_n = \max\{|f_n(x) - f(x)| : x \in [0,1]\} = d_{\max}(f_n, f).$$

Hence $(d_{\max}(f_n, f))$ is a null sequence, as claimed.

Conversely, let (f_n) be a sequence of functions in $C[0,1]$ that converges to $f \in C[0,1]$ with respect to $d_{\max}$. Then $(d_{\max}(f_n, f))$ is a null sequence. Again we can deduce from the Extreme Value Theorem that $f_n - f$ is bounded for each $n \in \mathbb{N}$ and that for each $n \in \mathbb{N}$,

$$\begin{aligned} M_n &= \sup\{|f_n(x) - f(x)| : x \in [0,1]\} \\ &= \max\{|f_n(x) - f(x)| : x \in [0,1]\} = d_{\max}(f_n, f), \end{aligned}$$

so (M_n) is also a null sequence. Hence $f_n \to f$ uniformly as $n \to \infty$. ■

Thus, if we are given a sequence of functions (f_n) in $C[0,1]$, we can use our knowledge of the properties of uniform convergence to determine whether it is convergent for $d_{\max}$, as follows.

Strategy: to determine convergence of sequences of functions in $C[0,1]$

1. Identify the possible limit function $f \in C[0,1]$ by finding the pointwise limit of (f_n).

   If there is no pointwise limit, then Lemma 4.5 implies that there is no uniform limit, and so Theorem 5.8 tells us that (f_n) is not convergent with respect to $d_{\max}$.

2. If the pointwise limit of (f_n) exists, determine whether it is continuous.

   If it is not continuous, then Lemma 4.5 tells us that the only possible uniform limit is not continuous, and Theorems 4.7 and 5.8 together then tell us that (f_n) is not convergent with respect to $d_{\max}$.

3. If f is the pointwise limit of (f_n) and is continuous, determine whether $(d_{\max}(f_n, f))$ is a null sequence.

   If it is, then by definition the sequence (f_n) converges to f for $d_{\max}$. Otherwise it does not converge in $C[0,1]$ for $d_{\max}$.

Worked Exercise 5.9

Consider the sequence of functions (f_n) where for each $n \in \mathbb{N}$, $f_n\colon [0,1] \to \mathbb{R}$ is given by

$$f_n(x) = \begin{cases} 1 - nx, & \text{if } 0 \leq x < \frac{1}{n}, \\ 0, & \text{if } \frac{1}{n} \leq x \leq 1. \end{cases}$$

(a) For each $x \in [0,1]$, find $\lim_{n\to\infty} f_n(x)$.

(b) Use the strategy to determine whether the sequence (f_n) converges in $C[0,1]$ for $d_{\max}$.

Solution

(a) Sketches of the graphs of f_n for the first few values of n (Figure 5.4) suggest that for each $x > 0$, $f_n(x)$ is eventually 0, whereas if $x = 0$, then $f_n(x) = 1$ for each $n \in \mathbb{N}$. Indeed, if $0 < x \leq 1$, then $f_n(x) = 0$ for $n > \frac{1}{x}$, and hence $f_n(x) \to 0$ as $n \to \infty$.

In addition $f_n(0) = 1$ for each $n \in \mathbb{N}$, and so $f_n(0) \to 1$ as $n \to \infty$.

Hence the pointwise limit of the sequence (f_n) is the function $f\colon [0,1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } 0 < x \leq 1. \end{cases}$$

(b) We use the strategy given above. Step 1 requires us to identify a possible limit function. By part (a), the only possible limit function for the sequence (f_n) is the pointwise limit function $f\colon [0,1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } 0 < x \leq 1. \end{cases}$$

But this function f is *not* continuous on $[0,1]$, and so is not in $C[0,1]$. Hence (f_n) does not converge in $C[0,1]$ for $d_{\max}$.

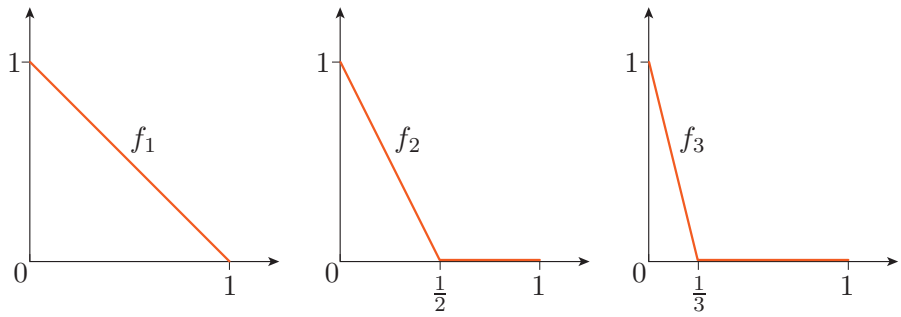


Figure 5.4 Sketches of f_1, f_2 and f_3 for Worked Exercise 5.9

Exercise 5.4

Consider the sequence of functions (f_n) where for each $n \in \mathbb{N}$, $f_n \colon [0,1] \to \mathbb{R}$ is given by

$$f_n(x) = \begin{cases} 2nx, & \text{if } 0 \leq x < \frac{1}{2n}, \\ 2 - 2nx, & \text{if } \frac{1}{2n} \leq x < \frac{1}{n}, \\ 0, & \text{if } \frac{1}{n} \leq x \leq 1. \end{cases}$$

(a) Sketch the graphs of f_1, f_2 and f_3.

(b) For each $x \in [0,1]$, find $\lim_{n\to\infty} f_n(x)$.

(c) Determine whether the sequence (f_n) converges in $C[0,1]$ for $d_{\max}$.

Learning outcomes

We list those topics on which we may set assessment questions to test your understanding of the chapter.

After working through this chapter you should be able to:

(a) Use a one–one map to transfer a metric from one set W to another set X.

(b) Given a metric on a set X, describe the induced metric on a subset of X.

(c) Construct a metric on the product of metric spaces.

(d) Define and use the Hamming distance.

(e) Define and use the Cantor metric.

(f) State the definition of a Lipschitz function, and determine whether a given function is Lipschitz.

(g) Explain why Lipschitz functions are continuous.

(h) Recognise that different metrics can give rise to the same continuous functions.

(i) Understand the relationship between continuity in a product space and its constituent spaces.

(j) Define pointwise convergence and uniform convergence for sequences of functions.

(k) Determine whether a given sequence of functions converges pointwise.

(l) Determine whether a given sequence of functions converges uniformly.

(m) Define and use the max metric on $C[0,1]$.

(n) Find the limit of a given sequence of functions in $C[0,1]$ for $d_{\max}$.

Solutions to exercises

Solution to Exercise 1.1

The distance function d is obtained from the pull-back of the Euclidean metric $d^{(1)}$ by the function $f(x) = \tan^{-1} x$. Hence, by Theorem 1.2, it is enough to show that f is one–one on $\mathbb{R}$ to conclude that d is a metric. Since $\tan^{-1}$ is a strictly increasing function (see Figure 1.2), it is indeed one–one. Thus d is a metric.

Solution to Exercise 1.2

The distance function d is obtained from the pull-back of the Euclidean metric $d^{(1)}$ by the function $f(x) = x^2$. However, f is not one–one, and so, by the remark after Theorem 1.2, d is not a metric.

(You can also check this directly since, for each $x \in \mathbb{R}$, $d(x, -x) = 0$, violating condition (M1) when $x \neq 0$.)

Solution to Exercise 1.3

Let (X, d) and (Y, e) be metric spaces and suppose that $f\colon X \to Y$ is given. Let $A \subseteq X$ and let d_A be the metric induced on A by d.

Suppose that $a \in A$ and f is (d, e)-continuous at a.

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in A that d_A-converges to a. (Thus $(d_A(a_n, a))$ is a null sequence.)

Then, from Theorem 1.5, (a_n) is a sequence in X that d-converges to a and so, since f is (d, e)-continuous at a, we conclude that the sequence $(f(a_n))$ e-converges to $f(a)$. But $f(a) = f|_A(a)$ and for each $n \in \mathbb{N}$, $f(a_n) = f|_A(a_n)$. Hence $(f|_A(a_n))$ e-converges to $f|_A(a)$ as $n \to \infty$.

In other words, for any choice of sequence (a_n) in A that d_A-converges to a, we have $(f|_A(a_n))$ e-converges to $f|_A(a)$ and so $f|_A$ is (d_A, e)-continuous at a.

Solution to Exercise 1.4

We find that

$$d^{(3)}(\mathbf{0}, (1, 0, 0)) = \sqrt{(1-0)^2 + (0-0)^2 + (0-0)^2} = 1,$$
$$d^{(3)}(\mathbf{0}, (0, 0, 1)) = \sqrt{(0-0)^2 + (0-0)^2 + (1-0)^2} = 1,$$

and so $(1, 0, 0)$ and $(0, 0, 1)$ are in S. Finally,

$$d_S^{(3)}((1, 0, 0), (0, 0, 1)) = \sqrt{(0-1)^2 + (0-0)^2 + (1-0)^2} = \sqrt{2}.$$

Solution to Exercise 1.5

(a) Let $\mathbf{a} = (a_1, a_2) \in S$ and $\mathbf{b} = (b_1, b_2) \in S$. In order to show that f is one–one, we need to show that if $f(\mathbf{a}) = f(\mathbf{b})$, then $\mathbf{a} = \mathbf{b}$.

So suppose that $f(\mathbf{a}) = f(\mathbf{b})$. Since, whatever the choice of $\mathbf{a} \in S$, $0 \leq f(\mathbf{a}) \leq 3$, we consider three cases.

$0 \leq f(\mathbf{a}) < 1$**:** In this case, $a_1 = 0$ and $f(\mathbf{a}) = 1 - a_2 = f(\mathbf{b})$ by our assumption. Now $a_2 \in [0, 1]$, and so $f(\mathbf{b}) \in [0, 1]$, which implies $b_1 = 0$ and $f(\mathbf{b}) = 1 - b_2$. That is, $1 - a_2 = 1 - b_2$ and $a_1 = 0 = b_1$, giving $\mathbf{a} = \mathbf{b}$.

$1 \leq f(\mathbf{a}) \leq 2$**:** In this case, $a_2 = 0$ and $f(\mathbf{a}) = a_1 + 1 = f(\mathbf{b})$ by our assumption. Now $f(\mathbf{a}) \in [1, 2]$, and so $f(\mathbf{b}) \in [1, 2]$, which implies $b_2 = 0$ and $f(\mathbf{b}) = b_1 + 1$. That is, $a_1 + 1 = b_1 + 1$ and $a_2 = b_2 = 0$, giving $\mathbf{a} = \mathbf{b}$.

$2 < f(\mathbf{a}) \leq 3$**:** In this case, $a_1 = 1$ and $f(\mathbf{a}) = a_2 + 2 = f(\mathbf{b})$ by our assumption. Now $a_2 + 2 \in (2, 3]$, and so $f(\mathbf{b}) \in (2, 3]$, which implies $b_1 = 1$ and $f(\mathbf{b}) = b_2 + 2$. That is, $a_2 + 2 = b_2 + 2$ and $a_1 = 1 = b_1$, giving $\mathbf{a} = \mathbf{b}$.

Thus in all three cases, $f(\mathbf{a}) = f(\mathbf{b})$ implies $\mathbf{a} = \mathbf{b}$. We conclude that f is one–one.

(b) Since f is a one–one mapping from S into $\mathbb{R}$ and d is the pull-back of the Euclidean metric on $\mathbb{R}$ by f, $d_f^{(1)}$, we deduce from Theorem 1.2 that d is a metric on S.

(c) We see that $f(\mathbf{a}) = f((0, 1)) = 0$ and $f(\mathbf{b}) = f((1, 1)) = 3$, so $d(\mathbf{a}, \mathbf{b}) = |0 - 3| = 3$. (Alternatively, since d measures distance 'along' S and S has length 3, it follows that $d(\mathbf{a}, \mathbf{b}) = 3$.)

In contrast, the Euclidean distance between the points $\mathbf{a}$ and $\mathbf{b}$ is 1 and so $d_S^{(2)}(\mathbf{a}, \mathbf{b}) = 1$.

Solution to Exercise 1.6

The proof that e_1 is a metric closely follows the proof given for the taxicab metric in Chapter 14.

(M1) For each $\mathbf{x}$, $\mathbf{y} \in X_1 \times X_2$, $d_1(x_1, y_1)$ and $d_2(x_2, y_2)$ are non-negative, since d_1 and d_2 are metrics, and hence so is their sum: thus $e_1(\mathbf{x}, \mathbf{y}) \geq 0$.

For each $\mathbf{x} \in X_1 \times X_2$,

$$e_1(\mathbf{x}, \mathbf{x}) = d_1(x_1, x_1) + d_2(x_2, x_2) = 0 + 0 = 0,$$

since d_1 and d_2 are metrics.

Conversely, suppose $\mathbf{x}, \mathbf{y} \in X_1 \times X_2$ are such that $e_1(\mathbf{x}, \mathbf{y}) = 0$. Then

$$0 = d_1(x_1, y_1) + d_2(x_2, y_2).$$

Since d_1 and d_2 are metrics, $d_1(x_1, y_1) \geq 0$ and $d_2(x_2, y_2) \geq 0$, so this equation implies that $d_1(x_1, y_1) = 0$ and $d_2(x_2, y_2) = 0$. Thus, since d_1 and d_2 are metrics, $x_1 = y_1$ and $x_2 = y_2$; that is, $\mathbf{x} = \mathbf{y}$. So e_1 satisfies (M1).

(M2) Let $\mathbf{x}, \mathbf{y} \in X_1 \times X_2$. Using property (M2) for the metrics d_1 and d_2, we have

$$\begin{aligned} e_1(\mathbf{y}, \mathbf{x}) &= d_1(y_1, x_1) + d_2(y_2, x_2) \\ &= d_1(x_1, y_1) + d_2(x_2, y_2) = e_1(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Thus e_1 satisfies (M2).

(M3) Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X_1 \times X_2$. Then, using property (M3) for the metrics d_1 and d_2,

$$\begin{aligned} e_1(\mathbf{x}, \mathbf{z}) &= d_1(x_1, z_1) + d_2(x_2, z_2) \\ &\leq (d_1(x_1, y_1) + d_1(y_1, z_1)) + (d_2(x_2, y_2) + d_2(y_2, z_2)) \\ &= (d_1(x_1, y_1) + d_2(x_2, y_2)) + (d_1(y_1, z_1) + d_2(y_2, z_2)) \\ &= e_1(\mathbf{x}, \mathbf{y}) + e_1(\mathbf{y}, \mathbf{z}). \end{aligned}$$

Thus e_1 satisfies (M3).

Therefore e_1 is a metric on $X_1 \times X_2$.

Solution to Exercise 1.7

$$\begin{aligned} e_1(\mathbf{0}, (2,1)) &= |2-0| + |1-0| = 3, \\ e_2(\mathbf{0}, (2,1)) &= \sqrt{(2-0)^2 + (1-0)^2} = \sqrt{5}, \\ e_\infty(\mathbf{0}, (2,1)) &= \max\{|2-0|, |1-0|\} = 2. \end{aligned}$$

Since these distances are all different, we conclude that the three metrics are distinct.

Solution to Exercise 1.8

(a) For the unit open ball for e_∞ centred at the origin, we have to find those $(x_1, x_2) \in \mathbb{R}^2$ satisfying

$$\max\{|x_1|, |x_2|\} < 1.$$

This is satisfied provided both $-1 < x_1 < 1$ and $-1 < x_2 < 1$, and the points satisfying these inequalities form the interior of the square shown in Figure 1.6.

(b) The unit closed balls centred at the origin are the regions shown in Figure 1.6 together with their boundary points, comprising the bounding diamond, circle and square.

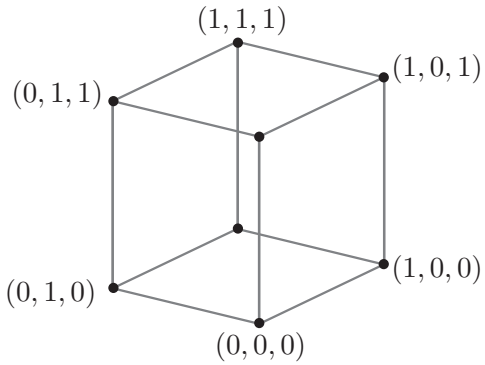
The unit spheres centred at the origin are simply the bounding diamond, circle and square of the regions shown in Figure 1.6.

Solution to Exercise 1.9

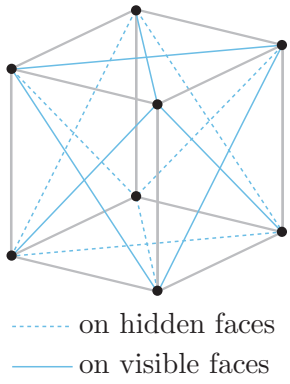
When $n = 3$,

$$A^n = A^3 = \{(0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$$

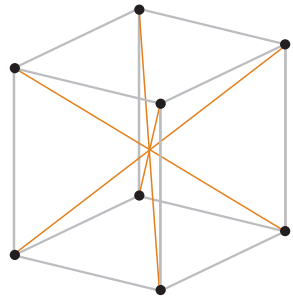
and this set of binary strings can be represented as the vertices of the usual unit cube in $\mathbb{R}^3$. Adjacent vertices are distance one apart and the distance between non-adjacent vertices is determined by the minimum number of edges required to join them, as shown in the figure.



(a) points that are distance 1 apart



(b) points that are distance 2 apart



(c) points that are distance 3 apart

Thus

$$B_{d_H}((0,0,0),r) = \begin{cases} \{(0,0,0)\}, & \text{if } 0 < r \leq 1, \\ \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}, & \text{if } 1 < r \leq 2, \\ A^3 - \{(1,1,1)\}, & \text{if } 2 < r \leq 3, \\ A^3, & \text{if } 3 < r. \end{cases}$$

Solution to Exercise 2.1

Here $\mathbf{a}$ and $\mathbf{e}$ first differ at their second term, so $d_{\mathbf{C}}(\mathbf{a},\mathbf{e}) = 2^{-2} = \frac{1}{4}$; $\mathbf{b}$ and $\mathbf{e}$ first differ at their third term, so $d_{\mathbf{C}}(\mathbf{b},\mathbf{e}) = 2^{-3} = \frac{1}{8}$; $\mathbf{c}$ and $\mathbf{e}$ first differ at their fourth term, so $d_{\mathbf{C}}(\mathbf{c},\mathbf{e}) = 2^{-4} = \frac{1}{16}$; and $\mathbf{d}$ and $\mathbf{e}$ first differ at their fifth term, so $d_{\mathbf{C}}(\mathbf{d},\mathbf{e}) = 2^{-5} = \frac{1}{32}$.

Solution to Exercise 2.2

We show that d satisfies conditions (M1)–(M3).

(M1) Since $|x| \geq 0$ whenever $x \in \mathbb{R}$, it follows that, for $(a_n), (b_n) \in \ell_0$,

$$d((a_n), (b_n)) = \sum_{n=1}^{\infty} |b_n - a_n| \geq 0.$$

Also,

$$d((a_n), (a_n)) = \sum_{n=1}^{\infty} |a_n - a_n| = 0,$$

and if $d((a_n), (b_n)) = 0$, then $\sum_{n=1}^{\infty} |b_n - a_n| = 0$, and so $a_n = b_n$ for each $n \in \mathbb{N}$. Hence in this case $(a_n) = (b_n)$.

Thus (M1) holds.

(M2) Let $(a_n), (b_n) \in \ell_0$. Then

$$\begin{aligned} d((a_n), (b_n)) &= \sum_{n=1}^{\infty} |b_n - a_n| \\ &= \sum_{n=1}^{\infty} |a_n - b_n| = d((b_n), (a_n)), \end{aligned}$$

and so (M2) holds.

(M3) Let $(a_n), (b_n), (c_n) \in \ell_0$. Then, for any $n \in \mathbb{N}$, the Triangle Inequality for the modulus function gives

$$|c_n - a_n| \leq |c_n - b_n| + |b_n - a_n|,$$

and so, using this inequality for each term of the sum,

$$\begin{aligned} \sum_{n=1}^{\infty} |c_n - a_n| &\leq \sum_{n=1}^{\infty} (|c_n - b_n| + |b_n - a_n|) \\ &= \sum_{n=1}^{\infty} |c_n - b_n| + \sum_{n=1}^{\infty} |b_n - a_n|. \end{aligned}$$

All of these sums have only a finite number of non-zero terms, so there is no need to call upon any of the theory of convergent series.

Thus

$$d((a_n), (c_n)) \leq d((a_n), (b_n)) + d((b_n), (c_n)),$$

as required.

Hence d is a metric on ℓ_0.

Solution to Exercise 2.3

Let $\mathbf{a} = (a_1, a_2, \ldots) \in \mathbf{C}$ and fix $n \in \mathbb{N}$. Recall that by definition

$$B_d[\mathbf{a}, 2^{-n}] = \{\mathbf{b} \in \mathbf{C} : d(\mathbf{a}, \mathbf{b}) \leq 2^{-n}\}.$$

If $n = 1$, then $2^{-n} = \frac{1}{2}$ and every element of $\mathbf{C}$ is within distance $\frac{1}{2}$ of $\mathbf{a}$. So in this case $B_d[\mathbf{a}, \frac{1}{2}] = \mathbf{C}$.

Suppose now that $n \geq 2$.

If $\mathbf{b} \in \mathbf{C}$ and there is an $i \in \{1, 2, \ldots, n-1\}$ for which $a_i \neq b_i$, then

$$d(\mathbf{a}, \mathbf{b}) \geq 2^{-i} \geq 2^{-n-1} > 2^{-n}$$

and so such a $\mathbf{b}$ is not in $B_d[\mathbf{a}, 2^{-n}]$.

Conversely, if $\mathbf{b} \in \mathbf{C}$ agrees with $\mathbf{a}$ for the first $n-1$ terms, then $\mathbf{b} = (a_1, a_2, \ldots, a_{n-1}, b_n, b_{n+1}, \ldots)$ and $\mathbf{b}$ cannot differ from $\mathbf{a}$ before the nth term. In which case, $d(\mathbf{a}, \mathbf{b}) \leq 2^{-n}$, implying that $\mathbf{b} \in B_d[\mathbf{a}, 2^{-n}]$.

We conclude that in this case

$$B_d[\mathbf{a}, 2^{-n}] = \{(a_1, a_2, a_3, \ldots, a_{n-1}, c_1, c_2, c_3, \ldots) : (c_i) \in \mathbf{C}\},$$

as required.

Solution to Exercise 3.1

We show that f is $(d_H, d_{\mathbf{C}})$-Lipschitz and use Theorem 3.2.

Let $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ in X. If $\mathbf{a} = \mathbf{b}$, then $d_H(\mathbf{a}, \mathbf{b}) = 0$ and $f(\mathbf{a}) = f(\mathbf{b})$ and so, in this case, we find

$$d_{\mathbf{C}}(f(\mathbf{a}), f(\mathbf{b})) = 0 = d_H(\mathbf{a}, \mathbf{b}).$$

If $\mathbf{a} \neq \mathbf{b}$, then they differ at between 1 and n places, so $1 \leq d_H(\mathbf{a}, \mathbf{b}) \leq n$. But it is always the case that $d_{\mathbf{C}}(\mathbf{x}, \mathbf{y}) \leq 2^{-1}$ no matter the choice of $\mathbf{x}$ and $\mathbf{y}$, so

$$d_{\mathbf{C}}(f(\mathbf{a}), f(\mathbf{b})) \leq 2^{-1} \leq 1 \leq d_H((a_1, a_2, a_3, \ldots, a_n), (b_1, b_2, b_3, \ldots, b_n)).$$

Hence f is $(d_H, d_{\mathbf{C}})$-Lipschitz (with Lipschitz constant 1) and we conclude by Theorem 3.2 that f is $(d_H, d_{\mathbf{C}})$-continuous on A^n.

Solution to Exercise 3.2

Let $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$. Then

$$\begin{aligned} e(\mathbf{x}, \mathbf{y}) &= |y_1 - x_1| + 2|y_2 - x_2| \\ &\leq 2|y_1 - x_1| + 2|y_2 - x_2| \\ &= 2(|y_1 - x_1| + |y_2 - x_2|) = 2e_1(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Similarly,

$$\begin{aligned} e_1(\mathbf{x}, \mathbf{y}) &= |y_1 - x_1| + |y_2 - x_2| \\ &\leq |y_1 - x_1| + 2|y_2 - x_2| \\ &= e(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Thus,

$$e_1(\mathbf{x}, \mathbf{y}) \leq e(\mathbf{x}, \mathbf{y}) \leq 2e_1(\mathbf{x}, \mathbf{y}).$$

Hence e and e_1 are equivalent.

Solution to Exercise 3.3

(a) Fix $f\colon X \to Y$ and $x \in X$.

Recall that f is (d_1, e)-continuous at x if, and only if, whenever (x_n) is a sequence in X for which $x_n \xrightarrow{d_1} x$ as $n \to \infty$, then $f(x_n) \xrightarrow{e} f(x)$ as $n \to \infty$.

But Proposition 3.7 tells us that $x_n \xrightarrow{d_1} x$ as $n \to \infty$ if, and only if, $x_n \xrightarrow{d_2} x$ as $n \to \infty$.

Hence f is (d_1, e)-continuous at x if, and only if, whenever (x_n) is a sequence in X for which $x_n \xrightarrow{d_2} x$ as $n \to \infty$, then $f(x_n) \xrightarrow{e} f(x)$ as $n \to \infty$, which is the same as f being (d_2, e)-continuous at x.

(b) Fix $g\colon Y \to X$ and $y \in Y$.

Recall that g is (e, d_1)-continuous at y if, and only if, whenever (y_n) is a sequence in Y for which $y_n \xrightarrow{e} y$ as $n \to \infty$, then $g(y_n) \xrightarrow{d_1} g(y)$ as $n \to \infty$.

But Proposition 3.7 tells us that for any sequence (x_n) in X, $x_n \xrightarrow{d_1} x$ as $n \to \infty$ if, and only if, $x_n \xrightarrow{d_2} x$ as $n \to \infty$, and so $g(y_n) \xrightarrow{d_1} g(y)$ if, and only if, $g(y_n) \xrightarrow{d_2} g(y)$ as $n \to \infty$.

Hence g is (e, d_1)-continuous at y if, and only if, whenever (y_n) is a sequence in Y for which $y_n \xrightarrow{e} y$ as $n \to \infty$, then $g(y_n) \xrightarrow{d_2} g(y)$ as $n \to \infty$, which is the same as g being (e, d_2)-continuous at y.

Solution to Exercise 3.4

(a) (i) $(0, (0, 0, 0, 0, 0, 0, \ldots))$.
(ii) $(0, (1, 0, 1, 0, 1, 0, \ldots))$.
(Notice the extra brackets.)

(b) Consider the composites of the projection functions with f, that is, $p_1 \circ f\colon \mathbf{C} \to \mathbb{R}$ and $p_2 \circ f\colon \mathbf{C} \to \mathbf{C}$. Now

$$(p_1 \circ f)(\mathbf{a}) = a_1 = g(\mathbf{a}),$$

where $g\colon \mathbf{C} \to \mathbb{R}$ is given by $g(\mathbf{a}) = a_1$. By assumption, g is $(d_{\mathbf{C}}, d^{(1)})$-continuous. Also,

$$(p_2 \circ f)(\mathbf{a}) = (a_2, a_3, a_4, \ldots) = \sigma(\mathbf{a}),$$

and we saw in Worked Exercise 2.6 that σ is $(d_{\mathbf{C}}, d_{\mathbf{C}})$-continuous.

Since both $p_1 \circ f$ and $p_2 \circ f$ are continuous, it follows from Theorem 3.12 that f is $(d_{\mathbf{C}}, e)$-continuous, where e is any one of e_1, e_2 or e_∞.

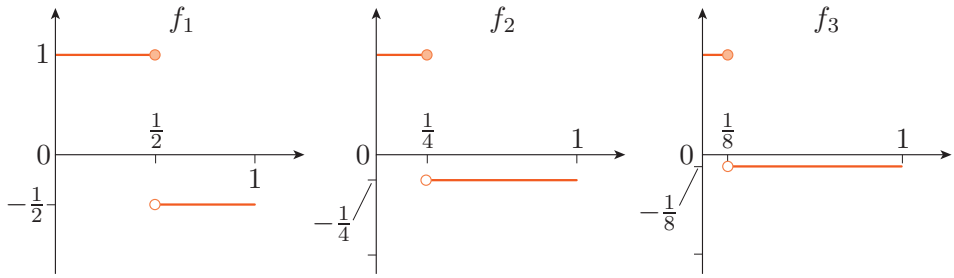
Solution to Exercise 3.5

The first component sequence converges for the Euclidean metric, since it is a basic null sequence.

However, the sequence $(\frac{1}{n^2})$ is not eventually constant and so, by the result of Exercise 3.2 of Chapter 14, does not converge for the discrete metric. Hence $((\frac{1}{n}, \frac{1}{n^2}))$ is divergent for this metric.

Solution to Exercise 4.1

Sketching the graphs of the functions f_n for $n = 1, 2$ and 3 suggests that, for each $x > 0$, $f_n(x) \to 0$ as $n \to \infty$.



Indeed, for $x > 0$, choose $N \in \mathbb{N}$ so that $\frac{1}{2^N} < x$; then, for each $n > N$, $f_n(x) = -\frac{1}{2^n}$. So, for $x > 0$, $f_n(x) \to 0$ as $n \to \infty$.

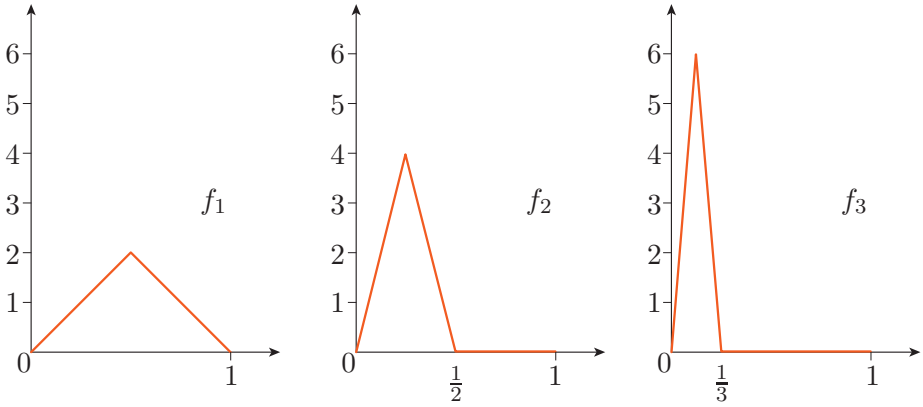
The behaviour at the point $x = 0$ is different: since $f_n(0) = 1$ for each $n \in \mathbb{N}$, $f_n(0) \to 1$ as $n \to \infty$.

So the pointwise limit of the sequence (f_n) is the (discontinuous) function $f : [0, \infty) \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } 0 < x < \infty. \end{cases}$$

Solution to Exercise 4.2

(a) The graphs of f_1, f_2 and f_3 are sketched below.



(b) We have $f_n(0) = 0$, for each $n \in \mathbb{N}$, and so $f_n(0) \to 0$ as $n \to \infty$.

If $x > 0$, then we can choose $N \in \mathbb{N}$ so that $\frac{1}{N} < x$. Then, for $n > N$, $f_n(x) = 0$. So $f_n(x) \to 0$ as $n \to \infty$.

Hence the pointwise limit function f of (f_n) is the zero function on $[0, 1]$. Thus $\int_0^1 f(x)\,dx = 0$.

(c) For each $n \in \mathbb{N}$, we find the integral of f_n by noting that the area representing the integral of f_n is a triangle with height $2n$ and width $\frac{1}{n}$. Thus

$$\int_0^1 f_n(x)\,dx = 1,$$

and so

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = 1.$$

(d) From part (b), $\int_0^1 \lim_{n\to\infty} f_n(x)\,dx = 0 \neq 1$ and so we conclude that the processes of integration and taking pointwise limits cannot, in general, be interchanged.

Solution to Exercise 4.3

We begin by identifying the pointwise limit of the sequence. For each $x \in [0, 1]$, $1 - (\frac{x}{n})^2 \to 1$ as $n \to \infty$, and so the pointwise limit of the sequence (f_n) is the constant function $f\colon [0, 1] \to \mathbb{R}$ given by $f(x) = 1$ for each $x \in [0, 1]$.

Thus, by Lemma 4.5, our only candidate for the uniform limit of the sequence (f_n) is the function f. To show that it is the uniform limit, we must show that the sequence (M_n), where M_n is given by

$$M_n = \sup\{|f_n(x) - f(x)| : x \in [0, 1]\},$$

is a null sequence.

To prove this, we observe that for each $x \in [0, 1]$ and each $n \in \mathbb{N}$,

$$\begin{aligned} 0 \leq |f_n(x) - f(x)| = |1 - (\tfrac{x}{n})^2 - 1| &= |\tfrac{x}{n}|^2 \\ &\leq \tfrac{1}{n^2}. \end{aligned}$$

Hence for each $n \in \mathbb{N}$,

$$0 \leq M_n = \sup\{|1 - (\tfrac{x}{n})^2 - 1| : x \in [0, 1]\} = \tfrac{1}{n^2}.$$

Since $(\frac{1}{n^2})$ is a basic null sequence, it follows that (M_n) is a null sequence.

Thus $f_n \to f$ uniformly as $n \to \infty$.

Solution to Exercise 4.4

For each $x \in \mathbb{R}$, $\frac{x}{n} \to 0$ as $n \to \infty$, and so the pointwise limit of (f_n) is the zero function given by $f(x) = 0$ for each $x \in \mathbb{R}$. Thus the only possible candidate for the uniform limit of the sequence (f_n) is f.

But $f_n - f$ is unbounded for each n: given any $M > 0$, $(f_n - f)(Mn) = f_n(Mn) - 0 = M$. Thus, for each $n \in \mathbb{N}$,

$$M_n = \sup\{|f_n(x) - f(x)| : x \in \mathbb{R}\} = +\infty$$

and so, in particular, (M_n) is not a real null sequence. Hence (f_n) does not converge uniformly to the zero function on $\mathbb{R}$.

Thus by Lemma 4.5, (f_n) is not uniformly convergent.

Solution to Exercise 4.5

(a) For each $x > 0$, we use the fact that $|\sin y| \leq 1$ for all $y \in \mathbb{R}$ to estimate

$$|f_n(x)| \leq \frac{1}{nx}.$$

Thus, for each $x > 0$, since $(\frac{1}{n})$ is a basic null sequence, we use the Multiple and Squeeze Rules to conclude that $f_n(x) \to 0$ as $n \to \infty$.

In contrast, we have $f_n(0) = 1$ for each $n \in \mathbb{N}$, and so $f_n(0) \to 1$ as $n \to \infty$.

Thus the sequence converges pointwise to the limit function $f\colon [0, \infty) \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x > 0. \end{cases}$$

This function is discontinuous at 0.

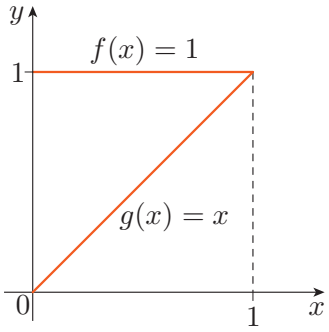
(b) Since

$$\lim_{y \to 0} \frac{\sin y}{y} = 1,$$

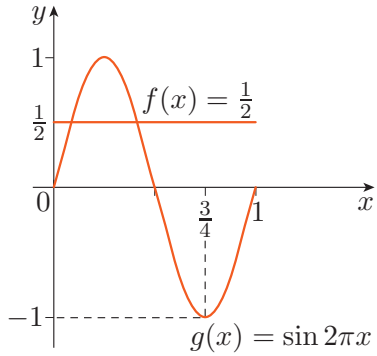
we deduce that for each $n \in \mathbb{N}$ the function f_n is continuous at $x = 0$. From Lemma 4.5 we know that the only possible uniform limit of the sequence (f_n) is the pointwise limit f. But if $f_n \rightrightarrows f$ as $n \to \infty$, then by Theorem 4.7 f would be continuous at 0. We conclude that the sequence (f_n) is not uniformly convergent.

Solution to Exercise 5.1

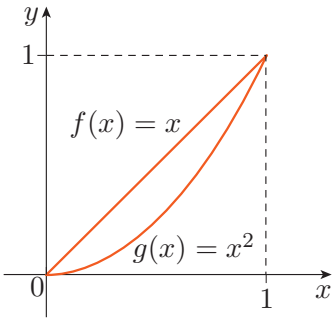
(a) Here the maximum value of $|g(x) - f(x)|$ is 1 (and is attained when $x = 0$); so $d_{\max}(f, g) = 1$.

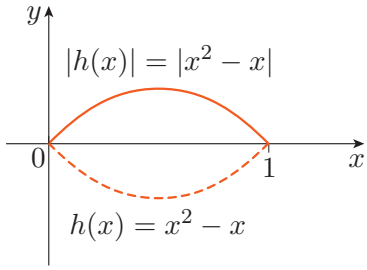


(b) In this case, the furthest $g(x)$ gets from $\frac{1}{2}$ is when $g(x) = -1$ (at $x = \frac{3}{4}$) and so $d_{\max}(f, g) = |(-1) - \frac{1}{2}| = \frac{3}{2}$.



(c) On $[0, 1]$, $g(x) \leq f(x)$ with equality at the endpoints only. Thus the maximum of $|g - f|$ over $[0, 1]$ occurs at the minimum of the function $h(x) = g(x) - f(x) = x^2 - x$ over this domain. This occurs at the point x_0 for which the derivative $h'(x) = 2x - 1$ vanishes, so that $x_0 = \frac{1}{2}$. Thus $d_{\max}(f, g) = |\frac{1}{4} - \frac{1}{2}| = \frac{1}{4}$.





Solution to Exercise 5.2

No, $d_{\min}$ does not define a metric on $C[0, 1]$, because it does not satisfy (M1). For example, if $f(x) = 0$ and $g(x) = x$ for $x \in [0, 1]$, then $d_{\min}(f, g) = 0$ but $f \neq g$.

Solution to Exercise 5.3

Let $h \in C[0,1]$ be given. Then $I_h = I \circ G_h$, where $G_h \colon C[0,1] \to C[0,1]$ is given by

$$G_h(f)(x) = h(x)f(x)$$

and $I \colon C[0,1] \to \mathbb{R}$ is given by

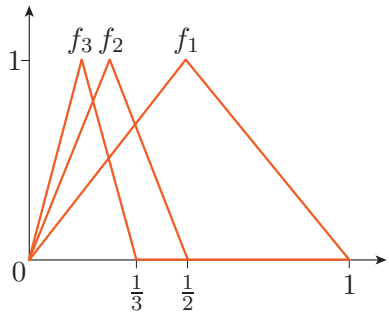
$$I(f) = \int_0^1 f(x)\,dx.$$

In Worked Exercises 5.5 and 5.6, we saw that G_h is $(d_{\max}, d_{\max})$-continuous on $C[0,1]$, and that I is $(d_{\max}, d^{(1)})$-continuous on $C[0,1]$.

Hence, by Proposition 4.4 of Chapter 14, with $X = Y = C[0,1], Z = \mathbb{R}$, $d_X = d_Y = d_{\max}$, $d_Z = d^{(1)}$, $f = G_h$ and $g = I$, we find that I_h is $(d_{\max}, d^{(1)})$-continuous on $C[0,1]$.

Solution to Exercise 5.4

(a) A sketch of the three graphs is shown below.



(b) For each $n \in \mathbb{N}$, $f_n(0) = 0$, and so $f_n(0) \to 0$ as $n \to \infty$.

If $0 < x \leq 1$, then $f_n(x) = 0$ for all $n > \frac{1}{x}$, and hence $f_n(x) \to 0$ as $n \to \infty$.

Thus, for each $x \in [0,1]$, $\lim_{n\to\infty} f_n(x) = 0$.

(c) We use the strategy. Step 1 of the strategy requires us to identify a possible limit. By part (b), the only possible limit for the sequence (f_n) is the pointwise limit function, that is, the zero function $f \colon [0,1] \to \mathbb{R}$ given by $f(x) = 0$ for each $x \in [0,1]$. This is a continuous function, and so we move to Step 3 of the strategy and calculate $d_{\max}(f_n, f)$, for $n \in \mathbb{N}$.

In fact, since $f_n(\frac{1}{2n}) = 1$ and $f_n(x) < 1$ for $x \neq \frac{1}{2n}$, we have, for each $n \in \mathbb{N}$,

$$d_{\max}(f_n, f) = \max\{|f_n(x) - 0| : x \in [0,1]\} = 1.$$

Thus the sequence $(d_{\max}(f_n, f))$ is not a null sequence, and so (f_n) does not converge in $C[0,1]$ for $d_{\max}$.